

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:01:52 ; Search time 188 seconds

(without alignments)  
539.875 Million cell updates/sec

Title: US-10-527-771-10

Perfect score: 1253

Sequence: 1 MSAAVVAVLALFSAENG.....DNLCDTRDAASVRHQCASP 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%  
Listing first 45 summaries

Database :

A\_Geneseq\_21.\*  
1: geneseqp19808.\*  
2: geneseqp19908.\*  
3: geneseqp20008.\*  
4: geneseqp20018.\*  
5: geneseqp20028.\*  
6: geneseqp20038.\*  
7: geneseqp20048.\*  
8: geneseqp20058.\*  
9: geneseqp20068.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1353	100.0	231	8	ADK13847
2	636.5	50.8	236	8	ADK13845
3	277	22.1	424	2	AAW04321
4	277	22.1	424	2	ADY37246
5	274	21.9	425	9	ADY37295
6	263	21.0	424	9	ADY37230
7	262.5	20.9	422	9	ADY37284
8	216.5	17.3	473	5	ABG31109
9	211	16.8	425	5	ABG31108
10	191	15.2	424	9	ADY37254
11	186.5	14.9	451	9	ADY37256
12	166	13.2	217	9	ADY37286
13	163	13.0	440	2	AA26786
14	160	12.8	216	9	ADY37296
15	154	12.3	207	9	ADY37297
16	153.5	12.3	218	2	AAW04322
17	152.5	12.2	218	2	AAW04323
18	147.5	11.8	218	8	ADY37248
19	137	10.9	227	8	ADY24015
20	137	10.9	229	8	ADY24017
21	137	10.9	238	8	AA123611
22	137	10.9	238	8	ADY24007
23	136.5	10.9	248	2	AA123609
24	136.5	10.9	248	8	ADY24005

25	135	10.8	200	9	ADY37250	Ady37250 Hookworm
26	132.5	10.6	220	2	AAW11485	AAW11485 O. volvuli
27	132.5	10.6	248	2	AAW11484	AAW11484 O. volvuli
28	131	10.5	454	2	AA26785	AA26785 Sequence
29	130.5	10.4	223	4	AAE13066	AAE13066 Onchocerca
30	130	10.4	232	8	ADY24013	ADY24013 Ancylostoma
31	126.5	10.1	253	4	AAE13067	AAE13067 Onchocerca
32	126	10.1	270	2	AA23608	AA23608 Canine ho
33	126	10.1	270	8	ADY24004	ADY24004 Ancylostoma
34	124.5	9.9	303	6	ABP96844	ABP96844 Human PR
35	124.5	9.9	371	6	ABU11626	ABU11626 Human KOD
36	124.5	9.9	395	8	ABG27603	ABG27603 Novel hum
37	124.5	9.9	395	8	ADY8739	ADY8739 Protein f
38	124.5	9.9	437	8	ABW84725	ABW84725 Human dia
39	124.5	9.9	437	8	ABW84726	ABW84726 Human dia
40	123.5	9.9	222	2	AAW37166	AAW37166 Haemonchu
41	123	9.8	247	4	ABE63677	ABE63677 Drosophila
42	123	9.8	415	7	ADY85552	ADY85552 Human NOV
43	122	9.7	255	5	AAE18962	AAE18962 Mouse tes
44	121	9.7	207	5	ABG31111	ABG31111 C. elegan
45	121	9.7	308	4	ABE64393	ABE64393 Drosophila

## ALIGNMENTS

RESULT 1	ADK13847	standard, protein, 231 AA.
ID	ADK13847	
AC	ADK13847	
XX		
DT	03-JUN-2004	(first entry)
DE	Ostertagia ostertagi immunogenic protein segid 10.	
XX		
DE	anti-parasitic; gene therapy; vaccine; nematode; worm;	
KW	Ostertagia ostertagi protein; immunogenic protein.	
XX		
OS	Ostertagia ostertagi.	
XX		
PN	US2004052817-A1.	
XX		
PD	18-MAR-2004.	
XX		
PF	13-SEP-2002; 2002US-00243319.	
XX		
PR	13-SEP-2002; 2002US-00243319.	
XX		
PA	(GRU/D) GRIJNDOP P.	
PA	(VERC/) VERCAUTEREN I.	
PA	(DMAE/) DE MAERE V.	
PA	(CLA/E/) CLAEREBOUT E.	
XX	(VERC/) VERCRUYSE J.	
PI	Geldhof P, Vercauteren I, De Maere V, Claerebout E, Vercruyse J;	
XX		
DR	WPI; 2004-247704/23.	
XX		
DR	N-PSDB; ADK13846.	
XX		
PT	New Ostertagia ostertagi proteins and nucleic acid sequences, useful for	
PT	diagnosing, preventing or treating infections caused by Ostertagia	
PT	ostertagi in animals.	
XX		
PS	Claim 16; SEQ ID NO 10; 36pp; English.	
XX		
CC	The invention describes a nucleic acid sequence (i) encoding an	
CC	Ostertagia ostertagi protein, or part of the nucleic acid sequence that	
CC	encodes an immunogenic fragment of the protein, where the nucleic acid	
CC	sequence or its part has at least 85% homology with the nucleic acid	
CC	sequence of the O. ostertagi gene having any of the 7 fully defined	
CC	sequences of 306-1761 base pairs, as given in the specification. The	
CC	composition and methods are useful for diagnosing, preventing or treating	

CC infections caused by *O. ostertagi* in animals. This is the amino acid  
 CC sequence of a novel immunogenic *Ostertagia ostertagi* protein.  
 XX  
 SO Sequence 231 AA;

Query Match 100.0%; Score 1253; DB 8; Length 231;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-118;  
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAVVVAVLALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIINMTVN 60  
 DB 1 MSAVVVAVLALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIINMTVN 60  
 QY 61 ADAVILGPAQNNYKYVDMDCNLEEVAAQOIAPCNDPLPINTSLAQNIAIRNLYFKQSEERTV 120  
 DB 61 ADAVILGPAQNNYKYVDMDCNLEEVAAQOIAPCNDPLPINTSLAQNIAIRNLYFKQSEERTV 120  
 QY 121 LQOVSWYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKICPAQNMVWSC 180  
 DB 121 LQOVSWYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKICPAQNMVWSC 180  
 QY 181 VYGSFKLAPNEVYIWOEGKACVCDARPDSCDNLCDTRDAASVRRHCCASP 231  
 DB 181 VYGSFKLAPNEVYIWOEGKACVCDARPDSCDNLCDTRDAASVRRHCCASP 231

## RESULT 2

ID ADK13845 standard; protein; 236 AA.

ADK13845;

03-JUN-2004 (first entry)

*Ostertagia ostertagi* immunogenic protein seqid 8.

anti-parasitic; gene therapy; vaccine; nematode; worm;

*Ostertagia ostertagi* protein; immunogenic protein.

*Ostertagia ostertagi*.

Key Location/Qualifiers

FT Misc-difference 37 /label= "LyS, Gln

FT Misc-difference 64 /note= "Encoded by MAA"

FT Misc-difference 111 /note= "Encoded by CAR"

FT Misc-difference 121 /note= "Encoded by CAR"

FT /label= "Asn, Thr

FT /note= "Encoded by AMT"

US2004052817-A1.

18-MAR-2004.

13-SEP-2002; 2002US-00243319.

13-SEP-2002; 2002US-00243319.

(GELD/) GELDHOF P.

(VERC/) VERCAUTEREN I.

(DMAE/) DE MAERE V.

(CLAE/) CLAEREBOU E.

(VERC/) VERCRUYSE J.

Geldhof P, Vercauteren I, De Maere V, Claerebout E, Vercruyse J;  
 WPI, 2004-247704/23.  
 DR N-PSDB; ADK13844.  
 XX New *Ostertagia ostertagi* proteins and nucleic acid sequences, useful for

PT diagnosing, preventing or treating infections caused by *Ostertagia*  
 PT *ostertagi* in animals.

PS Claim 15; SEQ ID NO 8; 36pp; English.

CC The invention describes a nucleic acid sequence (I) encoding an  
 CC *Ostertagia ostertagi* protein, or part of the nucleic acid sequence that  
 CC encodes an immunogenic fragment of the protein, where the nucleic acid  
 CC sequence or its part has at least 85% homology with the nucleic acid  
 CC sequence of the *O. ostertagi* gene having any of the 7 fully defined  
 CC sequences of 306-1761 base pairs, as given in the specification. The  
 CC composition and methods are useful for diagnosing, preventing or treating  
 CC infections caused by *O. ostertagi* in animals. This is the amino acid  
 CC sequence of a novel immunogenic *Ostertagia ostertagi* protein.

SO Sequence 236 AA;

Query Match 50.8%; Score 636.5; DB 8; Length 236;  
 Best Local Similarity 52.5%; Pred. No. 6.1e-56;  
 Matches 125; Conservative 35; Mismatches 67; Indels 11; Gaps 4;

QY 1 MSAVVVAV--LALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIINMT 58  
 DB 1 MSAVVVAV--LALPSYAEAGFCCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIINMT 58  
 QY 59 VNADAV---TLGPAQNNYKYVDMDCNLEEVAAQOIAPCNDPLPINTSLAQNIAIRNLYFKD 114  
 DB 59 VNADAV---TLGPAQNNYKYVDMDCNLEEVAAQOIAPCNDPLPINTSLAQNIAIRNLYFKD 114  
 QY 61 ---GAVQNRVYVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 172  
 DB 61 ---GAVQNRVYVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 172  
 QY 115 SEERTVLQOVSWYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 172  
 DB 115 SEERTVLQOVSWYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 172  
 QY 173 QONNVVSCVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 230  
 DB 173 QONNVVSCVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 230  
 QY 178 GTNVVWSVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 235  
 DB 178 GTNVVWSVYVWSASLGFMKGTKLDQFANQMAEPLANIANRNRKVGCAHKIC--PA 235

## RESULT 3

ID AAW04321 standard; protein; 424 AA.

AAW04321;

17-JAN-1997 (first entry)

*Ancylostoma secreted protein ASP-1 (pro-form)*.

*Ancylostoma secreted protein; ASP-1; hookworm; vaccine.*

*Ancylostoma caninum*.

Key Location/Qualifiers

FT Peptide 1..18 /label= "Sig\_peptide

FT Protein 19..424 /label= "Mat\_protein

WO9632479-A1.

17-OCT-1996.

10-APR-1996; 96WO-US004821.

10-APR-1995; 95US-00419414.

(UYVA ) UNIV YALE.

Hawdon JM, Hotez PJ, Jones BF;

WPI; 1996-477130/47.

DR N-PSDB; AAT38466.

PT Ancylostoma caninum secreted protein - useful as antigen for hookworm  
PT vaccine prodn.

Claim 2; Page 42-43; 66pp; English.

CC Ancylostoma secreted protein ASP-1 is secreted by canine hookworm larvae  
CC as they change from the free-living stage to the parasitic stage. The  
CC amino acid sequence of the ASP-1 pro-form (AA04321) was ded. from a  
CC cDNA clone (AA043466) obtd. from an Ancylostoma caninum L3 larvae cDNA  
CC library. ASP-1 represents a family of proteins (see also AA04322-23)  
CC that are highly immunogenic in experimental animals. Recombinant ASP  
CC proteins can be produced in a variety of hosts. They can be used in  
CC vaccines for hookworm, in the diagnosis of hookworm infection, or to  
CC raise antibodies

Sequence 424 AA;

Query Match 22.1%; Score 277; DB 2; Length 424;

Best Local Similarity 30.9%; Pred. No. 4.2e-19;

Matches 76; Conservative 35; Mismatches 97; Indels 38; Gaps 11;

QY 1 MSAVVAVVAVLAL-----PSYAEAGFCPCPSLSQSDSARQIFLDHNDVRNIALANGLIN 56  
DB 1 MSPPIVSVITFTIACDASPARDFGCSNS-GITDKRQAFLDHNNARRVAKG----- 54  
QY 57 WTVNADAVILGPAQNMVYKVDMDCNLEBVAAGQIAPCNDPLINTSLAONIARWLY---FK 113  
DB 55 -VEDNSGKLNPAKMYKLSWDCAMEQQLQDAIQSCPSAPAGIQVAGQNMVSSSGGFP 113  
QY 114 DSEETVLQGVSWYVNSASLGFMKGTAKDQ-----FANQWAEPLANIYANRNRKVGCA 166  
DB 114 DSVYK-IEQTLSGWMWSGAK---KNGVGPDKNYNGGLFA-----FSNMVYSETTKLGCA 163  
QY 167 HKICPAQNMVYSCYGSFKLAPNEVWQEGKACV---CDARPSFCNDLCTD-RDA 221  
DB 164 YKVGCTK--LAVSCYNGVGYITNQPMWETQACKTGADCTYKNSGCEDELCTGPDVP 221  
QY 222 SVRHQC 227  
DB 222 ETNQOC 227

RESULT 4

ID ADY37246 standard; protein; 424 AA.

XX ADY37246;

XX 05-MAY-2005 (first entry)

XX Hookworm antigen #9.

XX antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;

XX virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;

XX hookworm; immune response; HIV infection; tuberculosis; malaria; measles;

XX tetanus; diptheria; pertussis; polio; gene; de.

XX Ancylostoma caninum.

XX US200504232-A1.

XX 24-FEB-2005.

XX 16-APR-2004; 2004US-00825692.

XX 17-OCT-2001; 2001US-0329533P.

XX 23-NOV-2001; 2001US-0329533P.

XX 26-APR-2002; 2002US-0375404P.

XX 17-OCT-2002; 2002WO-US033106.

XX 26-SEP-2003; 2003US-0505848P.

XX (HOTEZ P.

XX (ASHC/) ASHCOW J.

PA (BDAM/) BDAMCHIAN M.

PA (ZHANG/) ZHANG B.

PA (WANG/) WANG Y.

PA (HARD/) HARDON J.

PA (LOIR/) LOIRAS A.

PA (WILL/) WILLIAMSON A.

PA (JONES/) JONES B.

PA (BETH/) BETHONY J.

PA (GOLD/) GOLD G.

PA (BORT/) BORTAZZI M E.

PA (MEND/) MENDEZ S.

PA Hotez P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hardon J,

PA Loukas A, Williamson A, Jones B, Bethony J, Gould G, Bortazzi ME,

PA Mendez S;

PA WPI; 2005-194935/20.

PA N-PSDB; ADY37245.

PA Composition useful for vaccinating or eliciting immune response against

PA hookworm in mammal, comprises copy of recombinant or synthetic antigen or

PA their fragments derived from hookworm, and carrier.

PA Disclosure; SEQ ID NO 18; 227pp; English.

XX The invention relates to a composition (I) comprising a copy of

XX recombinant or synthetic antigen or their fragments derived from hookworm

XX and a carrier. (I) is useful for vaccinating or eliciting an immune

XX response against hookworm in a mammal, reducing blood loss in a patient

XX infected with hookworm, reducing hookworm size or quantitative egg count

XX or hookworm burden in a patient infected with hookworm, or decreasing L3

XX migration across skin of a mammal. The above method further involves

XX chemically treating a hookworm-infected patient prior to the step of

XX administering. (M1) is also useful for vaccinating patient against

XX infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,

XX diptheria, pertussis or polio. This sequence corresponds to a protein of

XX the invention.

XX Sequence 424 AA;

XX Query Match 22.1%; Score 277; DB 9; Length 424;

XX Best Local Similarity 30.9%; Pred. No. 4.2e-19;

XX Matches 76; Conservative 35; Mismatches 97; Indels 38; Gaps 11;

XX 1 MSAVVAVVAVLAL-----PSYAEAGFCPCPSLSQSDSARQIFLDHNDVRNIALANGLIN 56

XX 1 MSPPIVSVITFTIACDASPARDFGCSNS-GITDKRQAFLDHNNARRVAKG----- 54

XX 57 WTVNADAVILGPAQNMVYKVDMDCNLEBVAAGQIAPCNDPLINTSLAONIARWLY---FK 113

XX 55 -VEDNSGKLNPAKMYKLSWDCAMEQQLQDAIQSCPSAPAGIQVAGQNMVSSSGGFP 113

XX 114 DSEETVLQGVSWYVNSASLGFMKGTAKDQ-----FANQWAEPLANIYANRNRKVGCA 166

XX 114 DSVYK-IEQTLSGWMWSGAK---KNGVGPDKNYNGGLFA-----FSNMVYSETTKLGCA 163

XX 167 HKICPAQNMVYSCYGSFKLAPNEVWQEGKACV---CDARPSFCNDLCTD-RDA 221

XX 164 YKVGCTK--LAVSCYNGVGYITNQPMWETQACKTGADCTYKNSGCEDELCTGPDVP 221

XX 222 SVRHQC 227

XX 222 ETNQOC 227

XX RESULT 5

XX ADY37295

XX ID ADY37295 standard; protein; 425 AA.

XX XX ADY37295;

XX 05-MAY-2005 (first entry)

DE Hookworm antigen #33.  
 XX anti-helminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquillizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio.  
 XX Ancylostoma duodenale.  
 OS  
 XX US2005042232-A1.  
 PN  
 XX 24-FEB-2005.  
 PD  
 XX 16-APR-2004; 2004US-00825692.  
 XX  
 XX 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US033106.  
 PR 26-SEP-2003; 2003US-0505848P.  
 XX  
 XX (HOTEZ P.  
 PA (ASHC/) ASHCOW J.  
 PA (BDAM/) BDMACHIAN M.  
 PA (ZHAN/) ZHAN B.  
 PA (WANG/) WANG Y.  
 PA (HAWD/) HAWDON J.  
 PA (LOUK/) LOUKAS A.  
 PA (WILL/) WILLIAMSON A.  
 PA (JONE/) JONES B.  
 PA (BETH/) BETHONY J.  
 PA (GOLD/) GOLD G.  
 PA (BOTT/) BOTTAZZI M E.  
 PA (MEND/) MENDEZ S.  
 XX  
 XX Hotez P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hawdon J;  
 PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME;  
 PI Mendez S;  
 XX WPI: 2005-194935/20.  
 DR  
 XX Composition useful for vaccinating or eliciting immune response against  
 PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
 PT their fragments derived from hookworm, and carrier.  
 XX  
 XX Disclosure; SEQ ID NO 67; 227bp; English.  
 PS  
 XX The invention relates to a composition (I) comprising a copy of  
 CC recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
 CC response against hookworm in a mammal, reducing blood loss in a patient  
 CC infected with hookworm, reducing hookworm size or quantitative egg count  
 CC or hookworm burden in a patient infected with hookworm, or decreasing IJ  
 CC migration across skin of a mammal. The above method further involves  
 CC chemically treating a hookworm-infected patient prior to the step of  
 CC administering. (M) is also useful for vaccinating patient against  
 CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
 CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
 CC the invention.  
 CC  
 XX  
 XX Sequence 425 AA:  
 SQ  
 Query Match 21.9%; Score 274; DB 9; Length 425;  
 Best Local Similarity 30.3%; Pred. No. 8.6e-19;  
 Matches 74; Conservative 37; Mismatches 97; Indels 36; Gaps 11;  
 OY 2 SAAVVAVLAL-----FSYAEAGFCPCNSISQSDSARQITLDPHNDVRRNTALG-----N 52  
 DB 3 SSVVVISVISTIAFCDSPARASFCGNN-GITDSRQATLDFHNNARRRVAAGVEDNKS 61  
 OY 53 GLINMTYNADAVILGPQNNYKVDWDCNLEBVAAGQIAFCNDPLPITSLAQNIAFLY- 111  
 DB 62 GKLN-----PAKNMYKLEWCKMEOQLADALQSCPGSGAGIQGFSQNVMSWSNS 110

OY 112 --FDSEETVLQGVSWYWSASLGFMKGTKLD-QFANQWAEPLANTANYRNRKVGCAHK 168  
 DB 111 GGPNSSEK-IEBTLGSMWGAK---NNGVGSNKKYGGGLYAFSNVVFSEETKIGCAVK 166  
 OY 169 ICPAQONMVVSCYVSGPKLAPNEVIMQEGKACV---CDARPDSFCCDNLCDT-RDAASV 223  
 DB 167 VCGTK-MATSCITNGIGYITNAPMWEQTGACKTGDCSTYKNSGCBDSICTKGADVPE 224  
 OY 224 RHQC 227  
 DB 225 NQC 228  
 RESULT 6  
 ADY37230  
 ID ADY37230 standard; protein; 424 AA.  
 XX  
 XX ADY37230;  
 AC  
 XX  
 DT 05-MAY-2005 (first entry)  
 DE  
 XX Hookworm antigen #1.  
 XX anti-helminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquillizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio; gene; ds.  
 XX Necator americanus.  
 OS  
 XX US2005042232-A1.  
 PN  
 XX 24-FEB-2005.  
 PD  
 XX 16-APR-2004; 2004US-00825692.  
 XX  
 XX 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US033106.  
 PR 26-SEP-2003; 2003US-0505848P.  
 XX  
 XX (HOTEZ P.  
 PA (ASHC/) ASHCOW J.  
 PA (BDAM/) BDMACHIAN M.  
 PA (ZHAN/) ZHAN B.  
 PA (WANG/) WANG Y.  
 PA (HAWD/) HAWDON J.  
 PA (LOUK/) LOUKAS A.  
 PA (WILL/) WILLIAMSON A.  
 PA (JONE/) JONES B.  
 PA (BETH/) BETHONY J.  
 PA (GOLD/) GOLD G.  
 PA (BOTT/) BOTTAZZI M E.  
 PA (MEND/) MENDEZ S.  
 XX  
 XX Hotez P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hawdon J;  
 PI Loukas A, Williamson A, Jones B, Bethony J, Goud G, Bottazzi ME;  
 PI Mendez S;  
 XX WPI: 2005-194935/20.  
 DR N-PSDB; ADY37229.  
 DR  
 XX  
 XX Composition useful for vaccinating or eliciting immune response against  
 PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
 PT their fragments derived from hookworm, and carrier.  
 XX  
 XX Disclosure; SEQ ID NO 2; 227bp; English.  
 PS  
 XX The invention relates to a composition (I) comprising a copy of  
 CC recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (I) is useful for vaccinating or eliciting an immune





SQ Sequence 425 AA;  
 Query Match 16.8%; Score 211; DB 5; Length 425;  
 Best Local Similarity 26.5%; Pred. No. 2.2e-12;  
 Matches 65; Conservative 36; Mismatches 76; Indels 68; Gaps 10;  
 QY 3 AAVVAVLALFEYAE-GECCPNSLSQSDARQIFLDFNDVRNIALG-----NGLIN 56  
 DB 2 AVLAVALLLACIERAAVQFGCSNT-KINDARKMEFDANDARRSMAGLBNKGL-- 58  
 QY 57 WTVNNAVILGPQNNYKVDMDCNLEBVAQAIAPC-----ND 94  
 DB 59 -----LSGKNVYELMDCEMRKAKQEMADGCPSSFGTFDPFWGQNTATYMGSIAD 109  
 QY 95 PLPINTSLAQNIRMLYFKDSEETVLQGVSWTVWSASIGPMKGTGLDQFANQMAEPLAN 154  
 DB 110 PLPY-ASMAVN-----GWSEIRTYGLDPPD--NKITSAMFRFAN 147  
 QY 155 IANYRNRYGCAHKICPAQNNVSCVYSPKLAPEVIMQEGKACVCDAR---PDSFC 210  
 DB 148 MANGKASAFGCAVALCAGK--LSINCITYNKIGYMTVAIYKGDACITSDAECTYSDSQ 205  
 QY 211 CNLNC 215  
 DB 206 KNGLC 210  
 RESULT 10  
 ID ADY37254 standard; protein; 424 AA.  
 AC ADY37254;  
 XX  
 DT 05-MAY-2005 (first entry)  
 XX  
 DE Hookworm antigen #13.  
 XX  
 KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio; gene; ds.  
 OS Ancylostoma caninum.  
 XX  
 PN US2005042232-A1.  
 XX  
 PD 24-FEB-2005.  
 XX  
 PF 16-APR-2004; 2004US-00825692.  
 XX  
 PR 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US031106.  
 PR 26-SEP-2003; 2003US-0505848P.  
 XX  
 PA (HOTE/) HOTEZ P.  
 PA (ASHC/) ASHCUM J.  
 PA (BDAM/) BDAMCHIAN M.  
 PA (ZHAN/) ZHAN B.  
 PA (WANG/) WANG Y.  
 PA (HARD/) HARDON J.  
 PA (LOUK/) LOUKAS A.  
 PA (WILL/) WILLIAMSON A.  
 PA (JONE/) JONES B.  
 PA (BETH/) BETHONY J.  
 PA (GOND/) GOND G.  
 PA (BOTT/) BOTTAZZI M E.  
 PA (MEND/) MENDEZ S.  
 XX  
 PI Hotez P, Ashcum J, Bdanchian M, Zhan B, Wang Y, Hawdon J,  
 PI Loukas A, Williamson A, Jones B, Bethony J, Gond G, Bottazzi ME,  
 PI Mendez S;

XX  
 DR WPI; 2005-194935/20.  
 DR N-PSDB; ADY37253.  
 XX  
 PT Composition useful for vaccinating or eliciting immune response against  
 PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
 PT their fragments derived from hookworm, and carrier.  
 PS Disclosure; SEQ ID NO 26, 227pp; English.  
 CC The invention relates to a composition (1) comprising a copy of  
 CC recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (1) is useful for vaccinating or eliciting an immune  
 CC response against hookworm in a mammal, reducing blood loss in a patient  
 CC infected with hookworm, reducing hookworm size or quantitative egg count  
 CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
 CC migration across skin of a mammal. The above method further involves  
 CC chemically treating a hookworm-infected patient prior to the step of  
 CC administering. (M) is also useful for vaccinating patient against  
 CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
 CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
 CC the invention.  
 XX  
 SQ Sequence 424 AA;  
 Query Match 15.2%; Score 191; DB 9; Length 424;  
 Best Local Similarity 29.5%; Pred. No. 2.3e-10;  
 Matches 62; Conservative 31; Mismatches 83; Indels 34; Gaps 11;  
 QY 23 CENLSQSDARQIFLDFNDVRNIALGNGLINVTNADAVILGPQNNYKVDMDCNLE 82  
 DB 21 CPGN-DLTDABRTLLTRVNSIRREIA--QGVAN--NVHGKLPAGKNIIYMRSCBLE 74  
 QY 83 EVA--AQQIAPCNDPLPINTSLAQNIRMLY-----FKDSEETVLQGVSWTVWSASL 133  
 DB 75 QAIDASQTF-CSASLBEPOKYGQNIQAVTPEIIRAPKNDLLEDAVXQ---WYLPVIY 130  
 QY 134 GPMKGTGLDQFANQMAE-----LANIANYRNRYGCAHKICPAQNNVSCVYSPKLAPE 189  
 DB 131 GGRDA-----ANKTDPRLTYFANLADKNTALGCHYAKCGPDRIVYSCMYNN-VVPD 183  
 QY 190 NEVIMQEGKACV---CDARPDSFCCDMLC 215  
 DB 184 NAVIYEPGACVKADCTTYPOSTCKDSLIC 213  
 RESULT 11  
 ID ADY37256 standard; protein; 451 AA.  
 AC ADY37256;  
 XX  
 DT 05-MAY-2005 (first entry)  
 XX  
 DE Hookworm antigen #14.  
 XX  
 KW antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KW virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
 KW hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KW tetanus; diphtheria; pertussis; polio; gene; ds.  
 OS Ancylostoma caninum.  
 XX  
 PN US2005042232-A1.  
 XX  
 PD 24-FEB-2005.  
 XX  
 PF 16-APR-2004; 2004US-00825692.  
 XX  
 PR 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002WO-US031106.  
 XX







XX Disclosure; SEQ ID NO 68; 227bp; English.

PS The invention relates to a composition (I) comprising a copy of  
 XX recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
 CC response against hookworm in a mammal, reducing blood loss in a patient  
 CC infected with hookworm, reducing hookworm size or quantitative egg count  
 CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
 CC migration across skin of a mammal. The above method further involves  
 CC chemically treating a hookworm-infected patient prior to the step of  
 CC administering. (M1) is also useful for vaccinating patient against  
 CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
 CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
 CC the invention.

XX Sequence 216 AA;

Query Match 12.8%; Score 160; DB 9; Length 216;  
 Best Local Similarity 27.4%; Pred. No. 1.4e-07;  
 Matches 68; Conservative 31; Mismatches 85; Indels 64; Gaps 15;

QY 5 VVAVVALLALPSYAEAG--FCCPNSLSQSDSARQITLDFHNDVRRVIALNGGLIMTVAD 62  
 DB 2 LVPVALLALLVAVAGNSMRCGN--GMTDEARQFLDVHNGYRSKVAQA-----KD 53  
 QY 63 AVITG-----PAQNMVKVMDCNLEBVAQAQIAPC---NDPLPINTSLAQNIAFWLYFQDS 115  
 DB 54 A--LGGNAPKAKMKMTYDCNVESTMDQAKCYFPAISHKGLGENTIMSTARQW----- 106  
 QY 116 EEEVTLQOVS--WYVWSASLGFPMKGTCLD-QFANQ-----WAEPLANINMYNR 161  
 DB 107 DKAQAQAQSDGMFALAKYGVGQENKLTQMLNKGVMIGHYTMWQWOB-----SY 157  
 QY 162 KYGCAHKICPAQQNMVSCVYGSPLKAP-----NVIWQEGKACVCDARPSFCCDNL-C 215  
 DB 158 KLGCVYEWCP-----SMTYGVCOYSPQGNMNSIYKGNPCT---QSDCGSNKAC 206

QY 216 DTRDAASV 223  
 DB 207 SSGEALCI 214

RESULT 15  
 ADY37297  
 ID ADY37297 standard; protein; 207 AA.

XX ADY37297;  
 AC 05-MAY-2005 (first entry)  
 DT  
 XX  
 DE Hookworm antigen #35.  
 XX  
 KM antihelminthic; anti-HIV; antitubercular; tuberculostatic; antimalarial;  
 KM virucide; tranquilizer; antibacterial; vaccine; recombinant antigen;  
 KM hookworm; immune response; HIV infection; tuberculosis; malaria; measles;  
 KM tetanus; diphtheria; pertussis; polio.  
 XX  
 OS Necator americanus.  
 XX  
 PN US200504223-A1.  
 PD 24-FEB-2005.  
 XX  
 PF 16-APR-2004; 2004US-00825692.  
 XX  
 PR 17-OCT-2001; 2001US-0329533P.  
 PR 23-NOV-2001; 2001US-0332007P.  
 PR 26-APR-2002; 2002US-0375404P.  
 PR 17-OCT-2002; 2002MO-US033106.  
 PR 26-SEP-2003; 2003US-0505848P.  
 XX  
 PA (HOTEL/ HOTEL P.

PA (ASHC/) ASHCOR J.  
 PA (BDAM/) BDAMCHIAN M.  
 PA (ZHAN/) ZHAN B.  
 PA (WANG/) WANG Y.  
 PA (HAWD/) HAWDON J.  
 PA (LOUK/) LOUKAS A.  
 PA (WILL/) WILLIAMSON A.  
 PA (JONE/) JONES B.  
 PA (BETH/) BETHONY J.  
 PA (GOUT/) GOUT G.  
 PA (BOTT/) BOTTAZZI M E.  
 PA (MENDE/) MENDEZ S.

PI Hotel P, Ashcom J, Bdamchian M, Zhan B, Wang Y, Hawdon J;  
 PI Loukas A, Williamson A, Jones B, Bethony J, Gout G, Bottazzi ME;  
 PI Mendez S;  
 DR WPI; 2005-194935/20.

XX Composition useful for vaccinating or eliciting immune response against  
 PT hookworm in mammal, comprises copy of recombinant or synthetic antigen or  
 PT their fragments derived from hookworm, and carrier.

XX Disclosure; SEQ ID NO 69; 227bp; English.

PS The invention relates to a composition (I) comprising a copy of  
 CC recombinant or synthetic antigen or their fragments derived from hookworm  
 CC and a carrier. (I) is useful for vaccinating or eliciting an immune  
 CC response against hookworm in a mammal, reducing blood loss in a patient  
 CC infected with hookworm, reducing hookworm size or quantitative egg count  
 CC or hookworm burden in a patient infected with hookworm, or decreasing L3  
 CC migration across skin of a mammal. The above method further involves  
 CC chemically treating a hookworm-infected patient prior to the step of  
 CC administering. (M1) is also useful for vaccinating patient against  
 CC infectious diseases e.g. HIV, tuberculosis, malaria, measles, tetanus,  
 CC diphtheria, pertussis or polio. This sequence corresponds to a protein of  
 CC the invention.

XX Sequence 207 AA;

Query Match 12.3%; Score 154; DB 9; Length 207;  
 Best Local Similarity 26.8%; Pred. No. 5.3e-07;  
 Matches 62; Conservative 35; Mismatches 84; Indels 50; Gaps 14;

QY 3 AA VVAVVALLALPSYAEAGFCCPNSLSQSDSARQITLDFHNDVRRVIALNGGLIMTVAD 62  
 DB 2 SSITCLVILSLIAVSKAG--CPDN--GMSSEARQFLDLNLSRVSVALGQADGAGNAP 58  
 QY 63 AVITGPAQNMVKVMDCNLEBVAQAQIAPC---NDPLPINTSLAQNIAFWLYFQDS 118  
 DB 59 -----KAAKMTMAVDCEVEKTMNNNAKQCVFKHSQPNQ--RKGIGENI---FMSSDSGKA 109  
 QY 119 TVLQOVS--WYVWSASLGFPMKGTCL--DQFAN-----QAEPLANINMYNRKRVGA 166  
 DB 110 KAAEQASXKWFELAEKGVQNLKTGGLFSRGVGHYTMWQWQETV-----KLGQY 160  
 QY 167 HKICPAQNM--VSCVYGSPLKAPNVIWQEGKACVCDARPSFCCDNLCD 216  
 DB 161 VEAC---SNMTCVVOYGPAGNMKGIDYKEGEC-----SKCEN-CD 199

Search completed: March 3, 2006, 19:05:33  
 Job time : 191 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:09:42 / Search time 47 Seconds  
(without alignments)  
406.342 Million cell updates/sec

Title: US-10-527-771-10

Perfect score: 1253  
Sequence: 1 MSAAVVAVLALFSYAENG.....DNLCPTDPAASVRHQCASP 231

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued Patents AA:\*

1: /cgn2\_6/prodata/1/1aa/5 COMB.pep:\*\n2: /cgn2\_6/prodata/1/1aa/6 COMB.pep:\*\n3: /cgn2\_6/prodata/1/1aa/8 COMB.pep:\*\n4: /cgn2\_6/prodata/1/1aa/ECTUS\_COMB.pep:\*\n5: /cgn2\_6/prodata/1/1aa/RB\_COMB.pep:\*\n6: /cgn2\_6/prodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	22.1	424	1	US-08-419-414-2
2	165	13.2	440	1	US-07-930-686-12
3	165	13.2	440	1	US-08-460-998-12
4	137	10.9	227	2	US-08-450-482B-125
5	137	10.9	227	2	US-08-151-064D-123
6	137	10.9	229	2	US-08-450-482B-127
7	137	10.9	229	2	US-08-151-064D-125
8	137	10.9	238	2	US-08-450-482B-117
9	137	10.9	238	2	US-08-151-064D-115
10	136.5	10.9	248	2	US-08-450-482B-115
11	136.5	10.9	248	2	US-08-151-064D-113
12	132.5	10.6	220	4	PCT-US96-07709-30
13	132.5	10.6	248	4	PCT-US96-07709-25
14	131	10.5	454	1	US-07-930-686-10
15	131	10.5	454	1	US-08-460-998-10
16	130.5	10.4	223	2	US-09-541-759-1
17	130	10.4	232	2	US-08-450-482B-123
18	130	10.4	232	2	US-08-151-064D-121
19	128.5	10.3	253	2	US-09-541-759-2
20	126	10.1	270	2	US-08-450-482B-114
21	126	10.1	270	2	US-08-151-064D-112
22	119	9.5	455	2	US-09-866-028-50
23	119	9.5	455	2	US-09-944-457-50
24	119	9.5	455	2	US-09-945-584-50
25	119	9.5	455	2	US-09-944-944-50
26	119	9.5	455	2	US-09-945-587-50
27	118	9.4	204	4	PCT-US96-07709-33

28	117.5	9.4	220	2	US-09-541-759-3	Sequence 3, Appl1
29	117	9.3	224	1	US-08-173-510B-87	Sequence 87, Appl
30	117	9.3	224	1	US-08-458-218-85	Sequence 85, Appl
31	117	9.3	224	1	US-08-450-497-87	Sequence 87, Appl
32	117	9.3	224	2	US-08-060-433C-37	Sequence 37, Appl
33	117	9.3	224	2	US-08-450-482B-87	Sequence 87, Appl
34	117	9.3	224	2	US-08-151-064D-85	Sequence 85, Appl
35	115.5	9.2	192	2	US-09-541-759-5	Sequence 5, Appl1
36	115.5	9.2	231	1	US-08-173-510B-86	Sequence 86, Appl
37	115.5	9.2	231	1	US-08-458-218-84	Sequence 84, Appl
38	115.5	9.2	231	1	US-08-450-497-86	Sequence 86, Appl
39	115.5	9.2	231	2	US-08-060-433C-36	Sequence 36, Appl
40	115.5	9.2	231	2	US-08-450-482B-86	Sequence 86, Appl
41	115.5	9.2	231	2	US-08-151-064D-84	Sequence 84, Appl
42	113	9.0	274	1	US-08-173-510B-83	Sequence 83, Appl
43	113	9.0	274	1	US-08-458-218-81	Sequence 81, Appl
44	113	9.0	274	1	US-08-450-497-83	Sequence 83, Appl
45	113	9.0	274	2	US-08-060-433C-33	Sequence 33, Appl

## ALIGNMENTS

RESULT 1  
US-08-419-414-2  
; Sequence 2, Application US/08419414  
; Patent No. 5753787  
; GENERAL INFORMATION:  
; APPLICANT: Hawdon, John M.  
; APPLICANT: Hotez, Peter J.  
; APPLICANT: Jones, Brian F.  
; TITLE OF INVENTION: Hookworm Vaccine  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; STREET: 1201 West Peachtree Street  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/419, 414  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: YU113  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8795  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 424 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Ancylostoma caninum  
; DEVELOPMENTAL STAGE: Larva  
; US-08-419-414-2  
Query Match 22.1%; Score 277; DB 1; Length 424;  
Best Local Similarity 30.9%; Pred. No. 2.2e-21;

	Matches	76; Conservative	35; Mismatches	97; Indels	38; Gaps	11;
QY		1	MSAAVYVAVLAL-----FSYAAGCCCNLSQSDSAQIFLDFENDVRBNIALGNGLIN	56		
Db		1	MFSPITVSIVITTLAFCDASPARDDGGCNS--GITDKDAFLDFENNAARRRYAKG-----	54		
QY		57	WTVAADAVILGPAONMYKVDMDCNLEEVAAQOIAFCNDPLPINTSLAONIAWLY--FK	113		
Db		55	-VEDNSDKLDPAAKMYKLSWDCNAEEQOLQDAIQSPSAFAGIQGVAAONVMSWSSSGGP	113		
QY		114	DSEETVYQQVSWYVVSASLGFPMKGTKLDPQ-----FANQWAEPLIANIANKRVGCA	166		
Db		114	DPSTVYIQTLTSGMWSGAK--KNGVGPDNKTNGGGLR-----FSNNVYSEITKLQGA	163		
QY		167	HKICPAQONMVSVCYGPKLAPNEVITQEGACV---CDARPDSPFCCDNICDT--RDA	221		
Db		164	YKVCSTK--LAVSCITNGVGYITTPMMWETGACXCTGADCSYKSGCEDLCTKGPDPV	221		
QY		222	SVRHQC	227		
Db		222	ETNQQC	227		

```

RESULT 2
US-07-930-686-12
; Sequence 12, Application US/07930686
; Patent No. 5525308
; GENERAL INFORMATION:
; APPLICANT: Sharp, Phillip J
; APPLICANT: Wagland, Barry M
; APPLICANT: Cobon, Gary S
; TITLE OF INVENTION: Nematode Vaccine
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley and Lardner
; STREET: suite 500, 1800 Diagonal Road
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 2213-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/930.686
; FILING DATE: 19921006
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK4486
; FILING DATE: 06-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU92/00040
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 440 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
Query Match 13.2%; Score 165; DB 1; Length 440;
Best Local Similarity 26.6%; Pred. NO.36-09;
Matches 59; Conservative 24; Mismatches 109; Indels 30; Gaps 9;

```

```

OY      15 SYABGPPCCPSLSGSDSAROIPLDFEHNDVBRNIALGCLIMTWAAADAVIILGPAONMYK  74
      : : : : :
Db      7 SCLRAAFPPISDNGMTDEIRQIFVJKNHBYRSIIIAKQ-----AKNKGGFAPKABRLK  61

OY      75 VDMDCNLEEEVAOQIAPC--NDPLPINTSLAQNIAWL-----YFKDSEETVLOQVW  126
      : : : : :
Db      62 VGYDEVEANNTAAVAKECFEHDPPEQRNRYWQNL--WMLGSGTNSYK--TESAKLSVOAW  117

OY      127 YWVASLSG----FMKGTIKDQFPAHQMAEPLANINRYBRKYGCAHKICGPAQONMVVSCV  181
      : : : : :
Db      118 YWELKMFQGPDENILITMEVFDRGVGHAYTO---VAMQSSDKIGCAVEMCPTM--TIVACE  171

OY      182 YGSPKTLAPNEVIMQEGKACVCDARPEDSFCCDNILCDTRDPAVY  223
      : : : : :
Db      172 YNPAGNRINHYIYDIDGPPCTTD--EDCQGTGCTGCSMDALCI  211

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Dh 7 SCLEAAFCPTSDNGMTDEIRQIFVDKHNBYRSIIANGQ-----AKNKLGGPAPBAABMLK 61
Qy 75 VMDDCNLEBAAQOAIAPC---NDPLEINTSLAQNARUL-----YFKOSEETVLQOOVSW 126
Dh 62 VGDYDCBEVANTAAVAKCEKFEHDPREQRYMGQN--RMLGGTYSK--TESAKLSVOAW 117
Qy 127 YWVSASLG-----FMKGTLDQDPANQMELPLIANYNRRKYGCAHKI CPAQNMVYSVCV 161
Dh 118 YMWLKNFGVPEDENILTWBEVPDRGVGHYTD---VAMQSSDKIGCAVEMCFPM--TLVACE 171
Qy 182 YGSPKLAPNEVIMQEGKACVCDARPPSPFCDDMLCPTRDAASY 223
Dh 172 YNPAGKRINHYYIDIDGPECTTD--EDCQCTGCTGCKEDALCI 211

```

```

US-08-450-482B-125
Sequence 125 Application US/08450482B
Patent No. 6818616
GENERAL INFORMATION:
APPLICANT: MATTHEW MOYLE ET AL.
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
NUMBER OF SEQUENCES: 132
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,482B
FILING DATE: 26-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/173,510
FILING DATE: 23-DEC-1993
APPLICATION NUMBER: 08/151,064
FILING DATE: 10-NOV-1993
APPLICATION NUMBER: 08/060,433
FILING DATE: 11-MAY-1993
APPLICATION NUMBER: 07/996,972
FILING DATE: 24-DEC-1992
APPLICATION NUMBER: 07/881,721
FILING DATE: 11-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: BIGGS, SUZANNE L.
REGISTRATION NUMBER: 30,158
REFERENCE/DOCKET NUMBER: 213/289
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 125:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 AMINO ACIDS
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
US-08-450-482B-125
Query Match 10.9%; Score 137; DB 2; Length 227;
Best Local Similarity 24.9%; Pred. No. 1.3e-06;
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps
22 CCGNSLSGSSANQRIPIIDFNDVRRNALDNGLIINTVNDAYILGPAQMMYKVDKCNL 8

```

Db 5 CKQNNSSMTNELRRRLRLHNGYRSLALGHVNI SESENEFFELVAHRAASRRLLIDYCD 64

Qy 82 EEVAAQO-IAPCNDPLPINTSLAONIARMLYFKDS---EEETVLQOVSWMYWSASLIGFM 136

Db 65 EGSAYESAIAKQSSNKSNSASBAYDENV---YVIDNTEYDEVDPALKAIS-SMTSQAENLT 119

Qy 137 KGTFLDQFPAQNAEPLANTANY---RARKVGCARHLCIQAQONMYVSCYCYGSPKLAPBEVI 193

Db 120 HAE--EGPIFYQNMDSVSPFANYAMDARREKLCGAVVTCDOGNTTHVHCYGPKAANKTEPI 177

Qy 194 WOEGRAC-----VCDARPPSPCCDNLCD 216

Db 178 YKVGVPSCNSCTBYTTKGEDEKVFCAH-DGVCVCYNLRD 213

```

RESULT 5
US-08-151-064D-123
; Sequence 123, Application US/08151064D
; Patent No. 6962795
; GENERAL INFORMATION:
; APPLICANT: MATTHEW MOYLE ET AL.
; TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
; NUMBER OF SEQUENCES: 130
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Word Perfect 5.1      IBM P.C. DOS 5.0
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/151,064D
; FILING DATE: 10-NOVEMBER-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/060,433
; FILING DATE: 11-MAY-1993
; APPLICATION NUMBER: 07/881,721
; FILING DATE: 11-MAY-1992
; APPLICATION NUMBER: 07/996,972
; FILING DATE: 24-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 203/226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULAR TYPE: PEPTIDE
; US-08-151-064D-123

Query Match          10.9%; Score 137; DB 2; Length 227,
Best Local Similarity 24.9%; Pred. No. 1,3e-06;
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8

QY      22 CCPNSISDSASRQIFLDFNHDVRNRNTALNGSLINMTVNADAVILGPAQNMYEVWDICNL 81
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB      5 CRGNNGSMYNELRRRFPLAHNGHSILALGHVINSSSESNFTFLYARASPMRLTDVDCOA 64

82 EEVAACQ-IAPCNDEPLINTSLAONIRMLTFKDS-----EEETVLDAQVSMTYWSASLGFM 136

```

Db 65 EGSAYESAIKQCSSNKKSSSAEYDENV---YVIDNTYEDEVDPAKKAIS-SWTSQAENLT 119  
QY 137 KGTKLQGFANQMAEPLANTANY---RNRKVGCAHKICPAQONWVSCVYSSPKLAPREVI 193  
Db 120 HAE--EGIPYQWNSVSDPANVAMDAKELGCAVVTCDQGNTHVCHYGPKAANKTEPI 177  
QY 194 WQEGKAC-----VCDARPDSPCCDNLCD 216  
Db 178 YKGVPCSNCTEYTRGDEKVFCHA-DEGVCVINLRD 213

RESULT 6  
US-08-450-482B-127  
Sequence 127, Application US/08450482B  
Patent No. 6818616  
GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Lyon & Lyon  
STREET: 633 West Filch Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,482B  
FILING DATE: 26-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/173,510  
FILING DATE: 23-DEC-1993  
APPLICATION NUMBER: 08/151,064  
FILING DATE: 10-NOV-1993  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/289  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 955-0440  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 127:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-450-482B-127

Query Match 10.9%; Score 137; DB 2; Length 229;  
Best Local Similarity 24.9%; Pred. No. 1.3e-06;  
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

QY 22 CCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIMTWVADAVILGPAQNMRYKVDMDCNL 81  
Db 7 CKQNGSMTELERRFLRLHNGYRSILALGHVNISSSENETFLYARASRMRLIDYDCDA 66  
QY 82 EEVAAQO-IAPCNDPLPITSLAQNIAKLYFKDS---EEETVLOQVSWYWSASLIGFM 136

Db 67 EGSAYESAIKQCSSNKKSSSAEYDENV---YVIDNTYEDEVDPAKKAIS-SWTSQAENLT 121  
QY 137 KGTKLQGFANQMAEPLANTANY---RNRKVGCAHKICPAQONWVSCVYSSPKLAPREVI 193  
Db 122 HAE--EGIPYQWNSVSDPANVAMDAKELGCAVVTCDQGNTHVCHYGPKAANKTEPI 179  
QY 194 WQEGKAC-----VCDARPDSPCCDNLCD 216  
Db 180 YKGVPCSNCTEYTRGDEKVFCHA-DEGVCVINLRD 215

RESULT 7  
US-08-151-064D-125  
Sequence 125, Application US/08151064D  
Patent No. 6962795  
GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Lyon & Lyon  
STREET: 633 West Filch Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/151,064D  
FILING DATE: 10-NOVEMBER-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 203/226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 125:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-151-064D-125

Query Match 10.9%; Score 137; DB 2; Length 229;  
Best Local Similarity 24.9%; Pred. No. 1.3e-06;  
Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

QY 22 CCPNSLSQSDSARQIFLDFHNDVRRNIALGNGLIMTWVADAVILGPAQNMRYKVDMDCNL 81  
Db 7 CKQNGSMTELERRFLRLHNGYRSILALGHVNISSSENETFLYARASRMRLIDYDCDA 66  
QY 82 EEVAAQO-IAPCNDPLPITSLAQNIAKLYFKDS---EEETVLOQVSWYWSASLIGFM 136  
Db 67 EGSAYESAIKQCSSNKKSSSAEYDENV---YVIDNTYEDEVDPAKKAIS-SWTSQAENLT 121

QY 137 KGTGLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQOMNVVSCVYSPKLAENYI 193  
DB 122 HAE--EGIPYQWNDVSDFANVAMDAAREKLGCAVVTCDQGNTHVCHGPRANKTEPI 179  
QY 194 MOEGKAC-----VCDARPDSPCCDNLCD 216  
DB 180 YKVGVCSCNCTEYTRGDBERKVFCHA-DEGVCVTLNLD 215

## RESULT 8

US-08-450-482B-117  
Sequence 117, Application US/08450482B  
Patent No. 6818616  
GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,482B  
FILING DATE: 26-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/173,510  
FILING DATE: 23-DEC-1993  
APPLICATION NUMBER: 08/151,064  
FILING DATE: 10-NOV-1993  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/289  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 117:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 238 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-450-482B-117

Query Match 10.9%; Score 137; DB 2; Length 238;  
Best Local Similarity 24.9%; Pred. No. 1.4e-06;

Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

QY 22 CCPNLSGSDSARQIFLDFHNDVRNIALGNGLIWTVNADAVIIGPAQOMTKVDMDCNL 81  
DB 16 CKQNNGSMTEHARRFLRHNGYRSIILAGHVNISEESNETFLVHARSRMRIILDYDCDA 75  
QY 82 BEVAAQO-IAPCNDPLPINTSLAONIARWLKFS-----EESTVLAQVSWVVSASLGFM 136  
DB 76 BGSAYESAIKQCSNKSXSSAEYDENV---YVIDNTYDEVDPAKALIS-SWTSQAFNLIT 130

QY 137 KGTGLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQOMNVVSCVYSPKLAENYI 193  
DB 131 HAE--EGIPYQWNDVSDFANVAMDAAREKLGCAVVTCDQGNTHVCHGPRANKTEPI 188  
QY 194 MOEGKAC-----VCDARPDSPCCDNLCD 216  
DB 189 YKVGVCSCNCTEYTRGDBERKVFCHA-DEGVCVTLNLD 224

## RESULT 9

US-08-151-064D-115  
Sequence 115, Application US/08151064D  
Patent No. 6962795  
GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/151,064D  
FILING DATE: 10-NOVEMBER-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 203/226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 115:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 238 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-151-064D-115

Query Match 10.9%; Score 137; DB 2; Length 238;  
Best Local Similarity 24.9%; Pred. No. 1.4e-06;

Matches 54; Conservative 30; Mismatches 103; Indels 30; Gaps 8;

QY 22 CCPNLSGSDSARQIFLDFHNDVRNIALGNGLIWTVNADAVIIGPAQOMTKVDMDCNL 81  
DB 16 CKQNNGSMTEHARRFLRHNGYRSIILAGHVNISEESNETFLVHARSRMRIILDYDCDA 75  
QY 82 BEVAAQO-IAPCNDPLPINTSLAONIARWLKFS-----EESTVLAQVSWVVSASLGFM 136  
DB 76 BGSAYESAIKQCSNKSXSSAEYDENV---YVIDNTYDEVDPAKALIS-SWTSQAFNLIT 130  
QY 137 KGTGLDQFANQWAEPLANIANY---RNRKVGCAHKICPAQOMNVVSCVYSPKLAENYI 193  
DB 131 HAE--EGIPYQWNDVSDFANVAMDAAREKLGCAVVTCDQGNTHVCHGPRANKTEPI 188

QY 194 WOEGKAC-----VCDARPSFCCDNLCD 216  
Db 189 YKGVPCSNCTEYTRGDEKVFCHA-DEGVCVIMLRD 224

RESULT 10  
US-08-450-482B-115  
; Sequence 115, Application US/08450482B  
; Patent No. 6818616

GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 132  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,482B  
FILING DATE: 26-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/173,510  
FILING DATE: 23-DEC-1993  
APPLICATION NUMBER: 08/151,064  
FILING DATE: 10-NOV-1993  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 213/289  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 115:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 248 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-450-482B-115

Query Match 10.9%; Score 136.5; DB 2; Length 248;  
Best Local Similarity 24.8%; Pred. No. 1.7e-06;  
Matches 60; Conservative 35; Mismatches 102; Indels 45; Gaps 11;

QY 2 SAAVVAVVLAALFSAVBAAGFCPPNSLSQ-----SDSARQIFLDFHNDVRNIALGGLIN 56  
Db 2 SYLVVLAALAGIAHANEHDPTCPONEVEMEKGFDDMRILKFLALHNGYSTKLALGH--VS 59  
QY 57 WTVNA-----DAVILGPAQNMKYKWDGNCLEEVAAQIAPCNDPLPINTSLAQNIAAML 110  
Db 60 ITRESSEEDYLDYLLVAPTSKRYLEYDCEAKSAVESAKKQCTTAFSSSTKTDENLOVIE 119  
QY 111 YFDSSEETVLQOVSVWVSASLGFNKGTYLDOFANQMAE-----PLANIANYRN----- 160  
Db 120 DDPDINHAAALKAIISWTEAFNL-----KKTGSGVYRSLTINSNFRANLAMDT 167

QY 161 RKVGAHRTICPAQNMVSVYGSPELAPNE--VWQEGKAC-VCDARPSFCC---DN 213  
Db 168 RKVGAHVPCPSGNTHTV-CHV--PKIVKGGKPIYSIGKPGCGNDYASKFCHADEG 224

QY 214 LC 215  
Db 225 VC 226

RESULT 11  
US-08-151-064D-113  
; Sequence 113, Application US/08151064D  
; Patent No. 6962795

GENERAL INFORMATION:  
APPLICANT: MATTHEW MOYLE ET AL.  
TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS  
NUMBER OF SEQUENCES: 130  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
STREET: Suite 4700  
CITY: Los Angeles  
STATE: California  
COUNTRY: U.S.A.  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
MEDIUM TYPE: Storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/151,064D  
FILING DATE: 10-NOVEMBER-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/060,433  
FILING DATE: 11-MAY-1993  
APPLICATION NUMBER: 07/881,721  
FILING DATE: 11-MAY-1992  
APPLICATION NUMBER: 07/996,972  
FILING DATE: 24-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: BIGGS, SUZANNE L.  
REGISTRATION NUMBER: 30,158  
REFERENCE/DOCKET NUMBER: 203/226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 113:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 248 AMINO ACIDS  
TYPE: AMINO ACID  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PEPTIDE  
US-08-151-064D-113

Query Match 10.9%; Score 136.5; DB 2; Length 248;  
Best Local Similarity 24.8%; Pred. No. 1.7e-06;  
Matches 60; Conservative 35; Mismatches 102; Indels 45; Gaps 11;

QY 2 SAAVVAVVLAALFSAVBAAGFCPPNSLSQ-----SDSARQIFLDFHNDVRNIALGGLIN 56  
Db 2 SYLVVLAALAGIAHANEHDPTCPONEVEMEKGFDDMRILKFLALHNGYSTKLALGH--VS 59  
QY 57 WTVNA-----DAVILGPAQNMKYKWDGNCLEEVAAQIAPCNDPLPINTSLAQNIAAML 110  
Db 60 ITRESSEEDYLDYLLVAPTSKRYLEYDCEAKSAVESAKKQCTTAFSSSTKTDENLOVIE 119  
QY 111 YFDSSEETVLQOVSVWVSASLGFNKGTYLDOFANQMAE-----PLANIANYRN----- 160



Db 120 DPEIDHAAALKAIIISWATERFNL-----NKTGEGVYRSILNINSFANLAMPDT 167  
Qy 161 -RRVGAHKICPAQNMVWSCVYSPKLAPE--VIMOSKAC-VCDARPDSPCC---DN 213  
Db 168 RERKVGCAVVKCRSGNTHV-CHY--PKIYKGEKPIYSIGKPCRGCDNYASKFPCADBG 224  
Qy 214 LC 215  
Db 225 VC 226

## RESULT 12

PCT-US96-07709-30  
; Sequence 30, Application PC/TUS9607709  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN  
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln St., Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/07709  
; FILING DATE: 23-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-30-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 220 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US96-07709-30

Query Match 10.6%; Score 132.5; DB 4; Length 220;  
Best Local Similarity 23.5%; Pred. No. 3.8e-06;  
Matches 57; Conservative 24; Mismatches 91; Indels 71; Gaps 9;  
Qy 4 AVVVAVLALFSAARGCCPNSLSQSDSARQIFLDHNDYRBNALANGLINMTV-NAD 62  
Db 9 AIVAVV-----TGNCPSGKLTALERRKI-----VGQNKYSDLINGKLKRN 52  
Qy 63 AVILGPAQNMVSCVYSPKLAPEVIAQIAPCNDPLPINTSLAQNIRMLYFKDS---EERT 119  
Db 53 GTYMRGKMLRLNDCKLE-----SSAQRMWANGCICFGHSPQQRG 94  
Qy 120 VLOQVSWTWASLASLFPKMGTKLDQFANQMAEPLANIYRNR-----161  
Db 95 VGENVYAYWSSVSGGLKKTAGTDAGKSWSELPTL--YENNPSSNMWTWKVAGQVLHPT 152  
Qy 162 -----KVGCAHKI-CPAQNMVWSCVYSPKLAPEVIMOSKACVCDARPDSPCCD 212  
Db 153 QMAKRTYKIGCGAVATQCDGGRTLIVICHYSPGNNVGEVIYQRANPC---KVDKDCYT 208  
Qy 213 NLG 215

Db 209 KKC 211

## RESULT 13

PCT-US96-07709-25  
; Sequence 25, Application PC/TUS9607709  
; GENERAL INFORMATION:  
; APPLICANT: Tripp, Cynthia A.  
; TITLE OF INVENTION: PARASITIC HELMINTH VENOM ALLERGEN  
; TITLE OF INVENTION: ANTIGEN 5-LIKE GENES AND PROTEINS  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross & McIntosh  
; STREET: 1700 Lincoln St., Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: U.S.A.  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/07709  
; FILING DATE: 23-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-30-PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 248 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US96-07709-25

Query Match 10.6%; Score 132.5; DB 4; Length 248;  
Best Local Similarity 23.5%; Pred. No. 4.5e-06;  
Matches 57; Conservative 24; Mismatches 91; Indels 71; Gaps 9;

Qy 4 AVVVAVLALFSAARGCCPNSLSQSDSARQIFLDHNDYRBNALANGLINMTV-NAD 62  
Db 37 AIVAVV-----TGNCPSGKLTALERRKI-----VGQNKYSDLINGKLKRN 80  
Qy 63 AVILGPAQNMVSCVYSPKLAPEVIAQIAPCNDPLPINTSLAQNIRMLYFKDS---EERT 119  
Db 81 GTYMRGKMLRLNDCKLE-----SSAQRMWANGCICFGHSPQQRG 122  
Qy 120 VLOQVSWTWASLASLFPKMGTKLDQFANQMAEPLANIYRNR-----161  
Db 123 VGENVYAYWSSVSGGLKKTAGTDAGKSWSELPTL--YENNPSSNMWTWKVAGQVLHPT 180  
Qy 162 -----KVGCAHKI-CPAQNMVWSCVYSPKLAPEVIMOSKACVCDARPDSPCCD 212  
Db 161 QMAKRTYKIGCGAVATQCDGGRTLIVICHYSPGNNVGEVIYQRANPC---KVDKDCYT 236  
Qy 213 NLG 215  
Db 237 KKC 239

RESULT 14  
US-07-930-686-10  
; Sequence 10, Application US/07930686  
; Patent No. 552508  
; GENERAL INFORMATION:

APPLICANT: Sharp, Phillip J  
APPLICANT: Wagland, Barry M  
APPLICANT: Cobon, Gary S  
TITLE OF INVENTION: Nematode Vaccine  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley and Lardner  
STREET: suite 500, 1800 Diagonal Road  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/930,686  
FILING DATE: 19921006  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK4486  
FILING DATE: 06-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU92/00040  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A  
REGISTRATION NUMBER: 29,768  
TELEPHONE: (703) 836-9300  
TELEFAX: (703) 683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-930-686-10

Query Match 10.5%; Score 131; DB 1; Length 454;  
Best Local Similarity 24.4%; Pred. No. 1.5e-05;  
Matches 54; Conservative 29; Mismatches 112; Indels 26; Gaps 9;  
QY 19 AGPCPNSLSQSDSARQIFLDFHNDVRRIALGNGLINWTVNAADVILGPAQNMVYVMD 78  
DB 252 AGSCPELNNGMTDEARKKFVVDKNEYRSLIAKQ-----AKGPGQFAPKABRMKVND 306  
QY 79 CNLEVAQAQIAPCNDPLPINTSLAQNIAW---LYFKDSE-----EETVLOQVSMTYWS 130  
DB 307 CDVEANAMWMSKTC--FGILT--AAMLKRWGNMNMSSKANNTKEAABAAVAWFGDL 362  
QY 131 ASLGFNKGTKLDQFANQMAEPLANTANTRNRKVCAHKICPAQNMVYSCVYSGPGLAPN 190  
DB 363 QKYGVENNVFTNVTYTLTKYSQLAWQSDRIGCVVPCWSSMTVV--CEYNPGGDLPG 421  
QY 191 EYIMQEGKACVCDARPDSPCCDNLCDTRDAASVRHQCASP 231  
DB 422 EALYDVGDPTKDA--DCQCPGCTC--SRDEG-----LCVAP 454

RESULT 15  
US-08-460-998-10  
Sequence 10, Application US/08460998  
Patent No. 5942413  
GENERAL INFORMATION:  
APPLICANT: Sharp, Phillip J  
APPLICANT: Wagland, Barry M  
APPLICANT: Cobon, Gary S  
TITLE OF INVENTION: Nematode Vaccine  
NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Foley and Lardner  
STREET: suite 500, 3000 K Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: United States of America  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,998  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/930,686  
FILING DATE: 06-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PK4486  
FILING DATE: 06-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/AU92/00040  
FILING DATE: 06-FEB-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A  
REGISTRATION NUMBER: 29,768  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-460-998-10

Query Match 10.5%; Score 131; DB 1; Length 454;  
Best Local Similarity 24.4%; Pred. No. 1.5e-05;  
Matches 54; Conservative 29; Mismatches 112; Indels 26; Gaps 9;  
QY 19 AGPCPNSLSQSDSARQIFLDFHNDVRRIALGNGLINWTVNAADVILGPAQNMVYVMD 78  
DB 252 AGSCPELNNGMTDEARKKFVVDKNEYRSLIAKQ-----AKGPGQFAPKABRMKVND 306  
QY 79 CNLEVAQAQIAPCNDPLPINTSLAQNIAW---LYFKDSE-----EETVLOQVSMTYWS 130  
DB 307 CDVEANAMWMSKTC--FGILT--AAMLKRWGNMNMSSKANNTKEAABAAVAWFGDL 362  
QY 131 ASLGFNKGTKLDQFANQMAEPLANTANTRNRKVCAHKICPAQNMVYSCVYSGPGLAPN 190  
DB 363 QKYGVENNVFTNVTYTLTKYSQLAWQSDRIGCVVPCWSSMTVV--CEYNPGGDLPG 421  
QY 191 EYIMQEGKACVCDARPDSPCCDNLCDTRDAASVRHQCASP 231  
DB 422 EALYDVGDPTKDA--DCQCPGCTC--SRDEG-----LCVAP 454

Search completed: March 3, 2006, 19:11:06  
Job time : 48 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using SW model

Run on: March 3, 2006, 19:10:28 / Search time 165 seconds  
(without alignments)  
584.961 Million cell updates/sec

Title: US-10-527-771-10

Perfect score: 1253

Sequence: 1 MSAVVAVVAVLALFSAVBAENG.....DNICDTRDAASVRHQCASP 231

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253	100.0	231	4	US-10-243-319C-10
2	636.5	50.8	236	4	US-10-243-319C-8
3	277	22.1	424	5	US-10-825-692-18
4	274	21.9	425	5	US-10-825-692-16
5	272.5	21.7	218	4	US-10-051-644B-5
6	263	21.0	424	5	US-10-825-692-2
7	262.5	20.9	422	5	US-10-825-692-56
8	216.5	17.3	473	4	US-10-051-644B-3
9	211	16.8	425	4	US-10-051-644B-1
10	191	15.2	424	5	US-10-825-692-26
11	186.5	14.9	451	5	US-10-825-692-28
12	166	13.2	217	5	US-10-825-692-58
13	160	12.8	216	5	US-10-825-692-68
14	154	12.3	207	5	US-10-825-692-69
15	147.5	11.8	218	5	US-10-825-692-20
16	135	10.8	200	5	US-10-825-692-22
17	124.5	9.9	395	5	US-10-450-763-57962
18	123	9.8	247	6	US-11-097-143-17823
19	123	9.8	415	4	US-10-309-290-84
20	122	9.7	205	4	US-10-051-644B-6
21	122	9.7	245	3	US-09-800-198-92
22	122	9.7	255	3	US-09-876-225-2
23	122	9.7	255	4	US-10-417-727-2
24	122	9.7	255	5	US-10-909-864B-2
25	121	9.7	207	4	US-10-051-644B-7
26	121	9.7	308	6	US-11-097-143-19971
27	121	9.7	446	3	US-09-726-643-47

28	121	9.7	446	3	US-09-790-264-2	Sequence 2, Appl
29	121	9.7	446	4	US-10-042-141-47	Sequence 47, Appl
30	121	9.7	446	4	US-10-269-353-2	Sequence 2, Appl
31	121	9.7	446	4	US-10-169-395-5	Sequence 5, Appl
32	121	9.7	446	5	US-10-919-272-47	Sequence 47, Appl
33	121	9.7	446	5	US-10-900-926-2	Sequence 2, Appl
34	119	9.5	455	3	US-09-866-028-50	Sequence 50, Appl
35	119	9.5	455	3	US-09-944-457-50	Sequence 50, Appl
36	119	9.5	455	3	US-09-944-457-50	Sequence 50, Appl
37	119	9.5	455	3	US-09-944-457-50	Sequence 50, Appl
38	119	9.5	455	3	US-09-945-587-50	Sequence 50, Appl
39	119	9.5	455	3	US-09-945-015-50	Sequence 50, Appl
40	119	9.5	455	3	US-09-944-396-50	Sequence 50, Appl
41	119	9.5	455	3	US-09-944-433-50	Sequence 50, Appl
42	119	9.5	455	3	US-09-943-762-50	Sequence 50, Appl
43	119	9.5	455	3	US-09-944-654-50	Sequence 50, Appl
44	119	9.5	455	3	US-09-943-851A-50	Sequence 50, Appl
45	119	9.5	455	3	US-09-944-413-50	Sequence 50, Appl

## ALIGNMENTS

```
RESULT 1
US-10-243-319C-10
; Sequence 10, Application US/10243319C
; Publication No. US20040052817A1
; GENERAL INFORMATION:
; APPLICANT: Universiteit Gent
; TITLE OF INVENTION: Ostertagia vaccine
; FILE REFERENCE: Unigent Ostertagia
; CURRENT APPLICATION NUMBER: US/10/243,319C
; CURRENT FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 10
; LENGTH: 231
; TYPE: PRY
; ORGANISM: Ostertagia ostertagi
US-10-243-319C-10

Query Match      100.0%; Score 1253; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 5.6e-119;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSAVVAVVAVLALFSAVBAENGFCCEPNSLSQSDSARQIFLDFHNDVRNRLALGGLINVTN 60
DB 1 MSAVVAVVAVLALFSAVBAENGFCCEPNSLSQSDSARQIFLDFHNDVRNRLALGGLINVTN 60
QY 61 ADAVILGPAOMVYKYDMDQNLBEVAAQOIAFPCNDPLPINTSLAONIARIWYKDSBEETV 120
DB 61 ADAVILGPAOMVYKYDMDQNLBEVAAQOIAFPCNDPLPINTSLAONIARIWYKDSBEETV 120
QY 121 LQVSWVWVVASLSGFMKTKLQDFANQWAEPLANIANRNRKVGCAHRCIPAQOMVWVSC 180
DB 121 LQVSWVWVVASLSGFMKTKLQDFANQWAEPLANIANRNRKVGCAHRCIPAQOMVWVSC 180
QY 181 VYGSFKLAPNRYIWOEGACVCDARPDSPCCDNLCTTRDAASVRHQCASP 231
DB 181 VYGSFKLAPNRYIWOEGACVCDARPDSPCCDNLCTTRDAASVRHQCASP 231

RESULT 2
US-10-243-319C-8
; Sequence 8, Application US/10243319C
; Publication No. US20040052817A1
; GENERAL INFORMATION:
; APPLICANT: Universiteit Gent
; TITLE OF INVENTION: Ostertagia vaccine
; FILE REFERENCE: Unigent Ostertagia
; CURRENT APPLICATION NUMBER: US/10/243,319C
; CURRENT FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 14
```

```
SOFTWARE: Patentin version 3.2
SEQ ID NO 8
LENGTH: 236
TYPE: PRT
ORGANISM: Osteragia osteragi
FEATURE:
NAME/KEY: misc feature
LOCATION: (37)-(37)
OTHER INFORMATION: The 'Xaa' at location 37 stands for Lys, or Gln.
FEATURE:
NAME/KEY: misc feature
LOCATION: (121)-(121)
OTHER INFORMATION: The 'Xaa' at location 121 stands for Asn, or Thr.
US-10-243-319C-8

Query Match
Best Local Similarity 50.8%; Score 636.5; DB 4; Length 236;
Pred. No. 3.2e-56;
Matches 125; Conservative 35; Mismatches 67; Indels 11; Gaps 4;

QY 1 MSAAVVAV--LALPSYAEAGFCCPNLSOSDSARQIFLDPHNDVRNIALGNGLINMT 58
DB 1 M0ALGIALYLVLTSTNTEAGCCPADLNTDEAKXIFLDHNOVRDRIAGASPLMLT 60
QY 59 VVADAV---ILGPAONMYKVDMDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLTPKD 114
DB 61 ---GAVQKRNVLGPAPKMYRMDMDCNLEKAKAKMTWCTTPIPIDTSPQNLQWLLFQN 117
QY 115 SEETVYLOQVSWYVWSASISGFMKGTKLDQFANQMAEPLANIYRNRKYGCAHKIC--PA 172
DB 118 SQEXEVLGTPTMSWYTAISLRNLQPTDEANITYWQIRPLSNINAWQNLKAGCAHKYCKEPT 177
QY 173 QONMYVSCYSGPKLAPNEVWQEGKACVCDARPSFCCDNLCPIRDASVSHOCAS 230
DB 178 GTNMVVSAYGSEVLQDNEVWMDKGTQWCMNAYPNSFCNNLDTIAATLKKQPKS 235

RESULT 3
US-10-825-692-18
Sequence 18, Application US/10825692
GENERAL INFORMATION:
APPLICANT: Hotiez, Peter
APPLICANT: Ashcom, James
APPLICANT: Bdamchian, Mahnaz
APPLICANT: Zhan, Bin
APPLICANT: Wang, Yan
APPLICANT: Hawdon, John
APPLICANT: Loukas, Alexander
APPLICANT: Williamson, Angela
APPLICANT: Jones, Brian
APPLICANT: Bethony, Jeffrey
APPLICANT: Goud, Gaddam
APPLICANT: Botazzi, Maria E.
APPLICANT: Mendez, Susana
TITLE OF INVENTION: Hookworm Vaccine
FILE REFERENCE: 03740007aa
CURRENT APPLICATION NUMBER: US/10/825,692
CURRENT FILING DATE: 2004-04-16
PRIOR APPLICATION NUMBER: US 60/329,533
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: US 60/332,007
PRIOR FILING DATE: 2001-11-23
PRIOR APPLICATION NUMBER: US 60/375,404
PRIOR FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: PCT US02/33106
PRIOR FILING DATE: 2002-10-17
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin version 3.2
SEQ ID NO 18
LENGTH: 424
TYPE: PRT
ORGANISM: Ancylostoma caninum
US-10-825-692-18
```

```
Query Match
Best Local Similarity 22.1%; Score 277; DB 5; Length 424;
Pred. No. 2.6e-19;
Matches 76; Conservative 35; Mismatches 97; Indels 38; Gaps 11;

QY 1 MSAVVAVVIALI---PSYAEAGFCCPNLSOSDSARQIFLDPHNDVRNIALGNGLIN 56
DB 1 MFSPIVSVIFTTAFCDASPARDFGCSNS-GITDKROAFLDPHNARRRVAG----- 54
QY 57 WTVNADAVILGPAPKMYKVDMDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLTY---FK 113
DB 55 -VEDNSGKLPAPKMYKVDMDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLTY---FK 113
QY 114 DSEETVYLOQVSWYVWSASISGFMKGTKLDQ-----FANQMAEPLANIYRNRKYGCA 166
DB 114 DSVK-IEQTLISGWSGAK---KNGVDPNKNYNGGJLFA-----FSNMVSETTKLGCA 163
QY 167 HKICPAQONMYVSCYSGPKLAPNEVWQEGKACV---CDARPSFCCDNLCPI-RDA 221
DB 164 YKVGCTK--LAVSCITNGVGYITNQPMWETGQACKTGADCTYINSGCEDGLCTKGPDPV 221
QY 222 SVRHQC 227
DB 222 ETNQC 227

RESULT 4
US-10-825-692-67
Sequence 67, Application US/10825692
GENERAL INFORMATION:
APPLICANT: Hotiez, Peter
APPLICANT: Ashcom, James
APPLICANT: Bdamchian, Mahnaz
APPLICANT: Zhan, Bin
APPLICANT: Wang, Yan
APPLICANT: Hawdon, John
APPLICANT: Loukas, Alexander
APPLICANT: Williamson, Angela
APPLICANT: Jones, Brian
APPLICANT: Bethony, Jeffrey
APPLICANT: Goud, Gaddam
APPLICANT: Botazzi, Maria E.
APPLICANT: Mendez, Susana
TITLE OF INVENTION: Hookworm Vaccine
FILE REFERENCE: 03740007aa
CURRENT APPLICATION NUMBER: US/10/825,692
CURRENT FILING DATE: 2004-04-16
PRIOR APPLICATION NUMBER: US 60/329,533
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: US 60/332,007
PRIOR FILING DATE: 2001-11-23
PRIOR APPLICATION NUMBER: US 60/375,404
PRIOR FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: PCT US02/33106
PRIOR FILING DATE: 2002-10-17
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patentin version 3.2
SEQ ID NO 67
LENGTH: 425
TYPE: PRT
ORGANISM: Ancylostoma duodenale
US-10-825-692-67

Query Match
Best Local Similarity 21.9%; Score 274; DB 5; Length 425;
Pred. No. 5.3e-19;
Matches 74; Conservative 37; Mismatches 97; Indels 36; Gaps 11;

QY 2 SAAVVAVVIALI---FSYAEAGFCCPNLSOSDSARQIFLDPHNDVRNIALG-----N 52
DB 3 SSVVAVSVISTAFCDASPARDFGCSNN-GITDSDROAFLDPHNARRRVAGVBDNKS 61
QY 53 GLINWTVNADAVILGPAPKMYKVDMDCNLEEVAAQOIAPCNDPLPINTSLAONIARWLTY- 111
```

```

Db      62 GRLN-----PAKMYKLEWCKRBOQLQDAIOSCPGSGACIOGFSQVMSWSNS 110
Qy      112 --FKDSEETVLQOVSWYVWSASLGFMTKTKLD--OPANQWAEPLANIYRNRKVCANR 168
Db      111 GGFPSSEK-ISTSLGWSGAK--NNGVSDNKTGCGGLAFSNNVPSSETTKLGCAYK 166
Qy      169 ICPAQNMVWSCVYGSFKLAPNEVIMQEGKACV---CDARPSPCCDNICDT-RDASV 223
Db      167 VCGTK--MATSCYINGYITNAPMRETGAQCTGADCSYKNSGCEBGLCTKGDVPEPT 224
Qy      224 RHQC 227
Db      225 NQOC 228

RESULT 5
US-10-051-644B-5
; Sequence 5, Application US/10051644B
; Publication No. US20030126625A1
; GENERAL INFORMATION:
; APPLICANT: Liu, et al.
; TITLE OF INVENTION: Screens and Assays for Agents Useful in Controlling
; TITLE OF INVENTION: Parasitic Nematodes
; FILE REFERENCE: 2002630-0012
; CURRENT APPLICATION NUMBER: US/10/051,644B
; CURRENT FILING DATE: 2003-03-04
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Clustal W
; OTHER INFORMATION: Alignment of VAR-1, VAR-2, and Selected Other
; OTHER INFORMATION: Nematode VA Proteins.
US-10-051-644B-5

Query Match      21.7%; Score 272.5; DB 4; Length 218;
Best Local Similarity 31.3%; Pred. No. 3.3e-19;
Matches 73; Conservative 35; Mismatches 88; Indels 37; Gaps 10;

Qy      1 MSAVVAVVLTAL-----PSYAEAGFCPPNSLSQSDSARQIFLDPHNDVRYNIALNGLIN 56
Db      1 MSPPIVSVFTIATCDASPARDFGCSNS-GITDKQAFLDFNNARRVAKG-----54

Qy      57 WTVNDAVILGPQNMVYVMDCNLEEVAAQOIAPCNDPLPINTSLAONIARMLY---FK 113
Db      55 -VEDNSGKLANPAKMYKLSMDCAHEQQLQDAIOSCPGSAFAGIQGVAAQNTWSWSSGGYP 113

Qy      114 DSEETVLQOVSWYVWSASLGFMTKTKLDQ-----FANQWAEPLANIYRNRKVCANR 166
Db      114 DSVSV-IBETLSGWSGAK--KNGVGPNNKYNGGLFPA-----FSNNVPSSETTKLGA 163

Qy      167 HKICPAQNMVWSCVYGSFKLAPNEVIMQEGKACV---CDARPSPCCDNIC 215
Db      164 YKVGCTK--LAVSCIYNGYITNAPMRETGAQCTGADCSYKNSGCEBGLC 214

RESULT 6
US-10-825-692-2
; Sequence 2, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela

```

```

; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria E.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: PCT US02/33106
; PRIOR FILING DATE: 2002-10-17
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Necator americanus
US-10-825-692-2

Query Match      21.0%; Score 263; DB 5; Length 424;
Best Local Similarity 30.5%; Pred. No. 6.9e-18;
Matches 73; Conservative 36; Mismatches 106; Indels 24; Gaps 9;

Qy      1 MSAVVAVVLTAL-----PSYAEAGFCPPNSLSQSDSARQIFLDPHNDVRYNIALNGLIN 56
Db      1 MSPPIVSVFTIATCDASPARDFGCSNS-GITDKQAFLDFNNARRVAKG-----54

Qy      57 WTVNDAVILGPQNMVYVMDCNLEEVAAQOIAPCNDPLPINTSLAONIARMLYFKDSE 116
Db      55 -LEDNSKRLMPAKMYKLSMDCAHEQQLQDAIOSCPGSAFAGIQGVAAQNTWSWSSGGYP 113

Qy      117 BETVLQO--VSWYVWSASLGFMTKTKLD--OPANQWAEPLANIYRNRKVCANR 173
Db      114 DSVKIBETLSGWSGAK--KNGVGPNNKYNGGLFPAFSNNVPSSETTKLGCAYVCGTK 170

Qy      174 QNMVWSCVYGSFKLAPNEVIMQEGKACV---CDARPSPCCDNICDT-RDASV 227
Db      171 --LAVSCIYNGYITNAPMRETGAQCTGADCSYKNSGCEBGLCTGPDVPEINQOC 227

RESULT 7
US-10-825-692-56
; Sequence 56, Application US/10825692
; Publication No. US20050042232A1
; GENERAL INFORMATION:
; APPLICANT: Hotez, Peter
; APPLICANT: Ashcom, James
; APPLICANT: Bdamchian, Mahnaz
; APPLICANT: Zhan, Bin
; APPLICANT: Wang, Yan
; APPLICANT: Hawdon, John
; APPLICANT: Loukas, Alexander
; APPLICANT: Williamson, Angela
; APPLICANT: Jones, Brian
; APPLICANT: Bethony, Jeffrey
; APPLICANT: Goud, Gaddam
; APPLICANT: Botazzi, Maria E.
; APPLICANT: Mendez, Susana
; TITLE OF INVENTION: Hookworm Vaccine
; FILE REFERENCE: 03740007aa
; CURRENT APPLICATION NUMBER: US/10/825,692
; CURRENT FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/329,533
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: US 60/332,007
; PRIOR FILING DATE: 2001-11-23
; PRIOR APPLICATION NUMBER: US 60/375,404

```



CURRENT APPLICATION NUMBER: US/10/825,692  
CURRENT FILING DATE: 2004-04-16  
PRIOR APPLICATION NUMBER: US 60/329,533  
PRIOR FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: US 60/332,007  
PRIOR FILING DATE: 2001-11-23  
PRIOR APPLICATION NUMBER: US 60/375,404  
PRIOR FILING DATE: 2002-04-26  
PRIOR APPLICATION NUMBER: PCT US02/33106  
PRIOR FILING DATE: 2002-10-17  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 26  
LENGTH: 424  
TYPE: PRT  
ORGANISM: Ancylostoma caninum  
US-10-825-692-26

Query Match 15.2%; Score 191; DB 5; Length 424;  
Best Local Similarity 29.5%; Pred. No. 1.5e-10;  
Matches 62; Conservative 31; Mismatches 83; Indels 34; Gaps 11;

23 CPNSLSQSDSARQIFLDFHNDVRNRLALGNGLINMTVNADAVILGPAQNMVYKVMDCNLE 82  
21 CPEN-DLTDERTLLTRVHNSIRREIA--QGVAN---NTHGKLPAGKNITRMRSCELE 74

83 EVA--AQAIPACNDPLPINTSLAQNIAWMLY-----FDSSEETVLOQVSWYWSASL 133  
75 QAALDSQTF-CGASLEEPKQNTQAYTTPSIIAPKDLLEDVAKQ---MYLVITY 130

134 GFMKGTLDQFANQMAEP---LANTANYNRRKYGCAHKICPAQNMVSVCSYSPKLAP 189  
131 GGRDA-----ANKFTDPRLYTPANLAYDKNTALGCHYAKCGPDRIVISCMYNN-VVPD 183

190 NEVIMBEGKACV---CDARPDSPCCDNLIC 215  
184 NAVIYEPGACVADADCTTYPOSTCDISLC 213

Db

RESULT 11  
US-10-825-692-28  
Sequence 28, Application US/10825692  
GENERAL INFORMATION:  
APPLICANT: Hotcz, Peter  
APPLICANT: Ashcom, James  
APPLICANT: Bdamchian, Mahnaz  
APPLICANT: Wang, Yan  
APPLICANT: Zhan, Bin  
APPLICANT: Hawdon, John  
APPLICANT: Loukas, Alexander  
APPLICANT: Williamson, Angela  
APPLICANT: Jones, Brian  
APPLICANT: Bethony, Jeffrey  
APPLICANT: Goud, Gaddam  
APPLICANT: Botazzi, Maria E.  
APPLICANT: Mendez, Susana  
TITLE OF INVENTION: Hookworm Vaccine  
FILE REFERENCE: 03740007aa  
CURRENT APPLICATION NUMBER: US/10/825,692  
CURRENT FILING DATE: 2004-04-16  
PRIOR APPLICATION NUMBER: US 60/329,533  
PRIOR FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: US 60/332,007  
PRIOR FILING DATE: 2001-11-23  
PRIOR APPLICATION NUMBER: US 60/375,404  
PRIOR FILING DATE: 2002-04-26  
PRIOR APPLICATION NUMBER: PCT US02/33106  
PRIOR FILING DATE: 2002-10-17  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 28  
LENGTH: 451

TYPE: PRT  
ORGANISM: Ancylostoma caninum  
US-10-825-692-28

Query Match 14.9%; Score 186.5; DB 5; Length 451;  
Best Local Similarity 27.1%; Pred. No. 4.6e-10;  
Matches 59; Conservative 39; Mismatches 95; Indels 25; Gaps 9;

7 VAVILALFSYAEAGFCPCNSLSQSDSARQIFLDFHNDVRNRLALGNGLINMTVNADAVI 65  
5 ILVVALILGIAHNTDFQCN--FKSTDTLREHYLKSINNLRKTIADG-----SAENKSGK 57

66 LGPAQNMVYKVMDCNLEEVAAQIAPCNDPLPINTSLAQNIAWMLYKDSSEET--VLQ 123  
58 CPQKNIYTLSDMCBELAQAQAVDQCKNVPEPAGYSQILKK---VKSTCPTVYKXQ 114

124 VSWYWSASLGFMKGTLDQFAN--QMAEPLANIANYNRRKYGCAHKICPAQNMVSVCSY 181  
115 IEAMWTKS---VKAGVDPNNKQGLDFAKLANGKATKICGAKKNC--NEQLYVACV 168

182 YGSPKLAPNEVIMBEGKAC---VCDARPDSPCCDNLIC 215  
169 INEPAPAVGMPIYEVGAGCNSKDDCTTYLQSKCSNKVC 206

Db

RESULT 12  
US-10-825-692-58  
Sequence 58, Application US/10825692  
GENERAL INFORMATION:  
APPLICANT: Hotcz, Peter  
APPLICANT: Ashcom, James  
APPLICANT: Bdamchian, Mahnaz  
APPLICANT: Wang, Yan  
APPLICANT: Zhan, Bin  
APPLICANT: Hawdon, John  
APPLICANT: Loukas, Alexander  
APPLICANT: Williamson, Angela  
APPLICANT: Jones, Brian  
APPLICANT: Bethony, Jeffrey  
APPLICANT: Goud, Gaddam  
APPLICANT: Botazzi, Maria E.  
APPLICANT: Mendez, Susana  
TITLE OF INVENTION: Hookworm Vaccine  
FILE REFERENCE: 03740007aa  
CURRENT APPLICATION NUMBER: US/10/825,692  
CURRENT FILING DATE: 2004-04-16  
PRIOR APPLICATION NUMBER: US 60/329,533  
PRIOR FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: US 60/332,007  
PRIOR FILING DATE: 2001-11-23  
PRIOR APPLICATION NUMBER: US 60/375,404  
PRIOR FILING DATE: 2002-04-26  
PRIOR APPLICATION NUMBER: PCT US02/33106  
PRIOR FILING DATE: 2002-10-17  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 58  
LENGTH: 217  
TYPE: PRT  
ORGANISM: Ancylostoma ceylanicum  
US-10-825-692-58

Query Match 13.2%; Score 166; DB 5; Length 217;  
Best Local Similarity 27.5%; Pred. No. 2.3e-08;  
Matches 63; Conservative 28; Mismatches 106; Indels 32; Gaps 8;

5 VVAVILALFSYAEAGFCPCNSLSQSDSARQIFLDFHNDVRNRLALGNGLINMTVNADAV 64  
3 VPLVILVALSVADANSVRGANGM--TDEARQKFLDMHNGYSQVAKGQAKALSGNAP-- 58

65 ILGPAQNMVYKVMDCNLEEVAAQIAPCNDPLPINTSLAQNIAWMLYFKDSSEETVLOQ 123

Db 59 ---KAARKKKVNYVCGVSTAMQNAKKCVFTHSHKGLGENI---WMTTAREMDKYKSAEQ 113  
QY 124 VSWYVWSASLGFPMKGTKLDQFANQWAEPLANIANYRNF-----KVGCAHKICPAQOMV 177  
Db 114 ASQGFSLAEYGVGFENKLTJWQWNRBNTQIGHYQWMDTYLGCYVWCS-----167  
QY 178 VSCYGSFKLAP-----NEVWQEGKACVCDARPSFCCDNLCDTRDA 221  
Db 168 -SMTYGVQYSPQGNMNSIIEKGNPCT---QSDCGSNARCTADRA 211

RESULT 13  
US-10-825-692-68  
; Sequence 68, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Loukas, John  
; APPLICANT: Hawdon, John  
; APPLICANT: Williams, Angela  
; APPLICANT: Jones, Brian  
; APPLICANT: Bethony, Jeffrey  
; APPLICANT: Goud, Gaddam  
; APPLICANT: Botazzi, Maria E.  
; APPLICANT: Mendez, Susana  
; TITLE OF INVENTION: Hookworm Vaccine  
; FILE REFERENCE: 03740007aa  
; CURRENT FILING DATE: 2004-04-16  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/329,533  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/332,007  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: US 60/375,404  
; PRIOR FILING DATE: 2002-04-26  
; PRIOR APPLICATION NUMBER: PCT US02/33106  
; PRIOR FILING DATE: 2002-10-17  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 68  
; LENGTH: 216  
; TYPE: PRT  
; ORGANISM: Ancylostoma duodenale  
US-10-825-692-68

Query Match 12.8%; Score 160; DB 5; Length 216;  
Best Local Similarity 27.4%; Pred. No. 9.1e-08;  
Matches 68; Conservative 31; Mismatches 85; Indels 64; Gaps 15;

QY 5 VVVAVVLLAFSYVAG--FCCPSLSOSDSARQIFLDFHNDVRNRIALGNGLINWTVND 62  
Db 2 LVFVALLALAVAVEGSMKGN--GMTDEARQEPFDVNHGYSKAKGA-----KD 53  
QY 63 AVILG-----PAOMYKVDMDCNLEEVAAQOIAPC---NDPLPINTSLAONTIARWLYFKDS 115  
Db 54 A--LGGNAPKAKMKKTIYCNVESTMDAKKCVFPHSHKGLGENIYMSSTARQM-----106  
QY 116 EEFVLQOVS--WYVWSASLGFPMKGTLD--QFANQ-----WAEPLANIANYRNR 161  
Db 107 DKAEAAQASDGMFAELAKYGVQENKLTJWQWNRBNTQIGHYQWMDTYLGCYVWCS-----SY 157  
QY 162 KVGCAHKICPAQOMVSCYGSFKLAP-----NEVWQEGKACVCDARPSFCCDNLCDTRDA 221  
Db 158 KLGCVVWCP-----SMTYGVQYSPQGNMNSIIEKGNPCT---QSDCGSNARCTADRA 206  
QY 216 DTRDAASV 223  
Db 207 SSGEALCT 214

RESULT 14  
US-10-825-692-69  
; Sequence 69, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Loukas, John  
; APPLICANT: Hawdon, John  
; APPLICANT: Williams, Angela  
; APPLICANT: Jones, Brian  
; APPLICANT: Bethony, Jeffrey  
; APPLICANT: Goud, Gaddam  
; APPLICANT: Botazzi, Maria E.  
; APPLICANT: Mendez, Susana  
; TITLE OF INVENTION: Hookworm Vaccine  
; FILE REFERENCE: 03740007aa  
; CURRENT FILING DATE: 2004-04-16  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/329,533  
; PRIOR FILING DATE: 2001-10-17  
; PRIOR APPLICATION NUMBER: US 60/332,007  
; PRIOR FILING DATE: 2001-11-23  
; PRIOR APPLICATION NUMBER: US 60/375,404  
; PRIOR FILING DATE: 2002-04-26  
; PRIOR APPLICATION NUMBER: PCT US02/33106  
; PRIOR FILING DATE: 2002-10-17  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 69  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Necator americanus  
US-10-825-692-69

Query Match 12.3%; Score 154; DB 5; Length 207;  
Best Local Similarity 26.8%; Pred. No. 3.5e-07;  
Matches 62; Conservative 35; Mismatches 84; Indels 50; Gaps 14;

QY 3 AAVVAVVLLAFSYVAGGFCPSLSOSDSARQIFLDFHNDVRNRIALGNGLINWTVND 62  
Db 2 SSTCLVTLSTAIYKAG--CPDN--GMSERARQKLELHNSRSSVALGQADGAGNAP 58  
QY 63 AVILGPAQOMYKVDMDCNLEEVAAQOIAPC---NDPLPINTSLAONTIARWLYFKDS 118  
Db 59 -----KAAMKTMAYDCEVEKTAAMNNAKQCVFKHSQPNQ--RKGLGENI---FMSSDSGKA 109  
QY 119 TVLQOVS--WYVWSASLGFPMKGTLD--QFANQ-----QWAEPLANIANYRNRKRVGA 166  
Db 110 KAEQASKAMFELERKGVQNLKLTGGLFSRGVGHYQWMDTYLGCYVWCS-----KLGCT 160  
QY 167 HKICPAQOM--VSCYGSFKLAPNEVWQEGKACVCDARPSFCCDNLCD 216  
Db 161 VZAC--SNMCTVVCQYGPAGMMGKDIYEKEBPC-----SKEN--CD 199

RESULT 15  
US-10-825-692-20  
; Sequence 20, Application US/10825692  
; Publication No. US20050042232A1  
; GENERAL INFORMATION:  
; APPLICANT: Hotez, Peter  
; APPLICANT: Ashcom, James  
; APPLICANT: Bdamchian, Mahnaz  
; APPLICANT: Zhan, Bin  
; APPLICANT: Wang, Yan  
; APPLICANT: Loukas, John  
; APPLICANT: Williams, Angela





*This Page Blank (uspto)*

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:11:22 ; Search time 21 Seconds

(without alignments)  
220.005 Million cell updates/sec

Title: US-10-527-771-10

Sequence: 1 MSAVVAVVLTALFSYAENG.....DNLCITRDAAVRRHCCASP 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 135346 seqs, 20000420 residues

Total number of hits satisfying chosen parameters: 135346

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppaa/BCT\_NEW\_PUB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1253	100.0	231	6	US-10-527-771-10
2	636.5	50.8	236	6	US-10-527-771-2
3	130.5	10.4	223	7	US-11-153-222A-1
4	128.5	10.3	253	7	US-11-153-222A-2
5	119	9.5	243	6	US-10-515-868-5
6	117.5	9.4	220	7	US-11-153-222A-3
7	111.5	8.9	258	7	US-11-090-439-38
8	106.5	8.5	245	6	US-10-515-868-6
9	103.5	8.3	500	7	US-11-067-573-2
10	99	7.9	243	6	US-10-515-868-4
11	99	7.9	243	6	US-10-506-443A-41
12	97.5	7.8	277	6	US-10-527-500-23
13	92.5	7.4	497	6	US-10-453-372-100
14	92.5	7.4	497	6	US-10-453-372-102
15	91.5	7.3	475	6	US-10-453-372-98
16	91.5	7.3	477	6	US-10-453-372-90
17	91.5	7.3	497	6	US-10-453-372-92
18	90.5	7.2	501	6	US-10-453-372-96
19	89.5	7.1	249	6	US-10-515-868-1
20	86.5	6.9	176	7	US-11-031-206-86
21	86	6.9	498	6	US-10-453-372-94
22	84.5	6.7	453	6	US-10-453-372-82
23	84.5	6.7	453	6	US-10-453-372-108
24	84.5	6.7	453	6	US-10-453-372-110
25	84.5	6.7	453	6	US-10-453-372-112

26	84	6.7	370	6	US-10-641-678-35	Sequence 35, Appl
27	84	6.7	459	6	US-10-641-678-72	Sequence 72, Appl
28	83.5	6.7	453	6	US-10-453-372-104	Sequence 104, Appl
29	82.5	6.6	236	6	US-10-527-500-65	Sequence 65, Appl
30	81.5	6.5	4555	6	US-10-995-561-556	Sequence 556, Appl
31	81	6.5	1332	7	US-11-242-459-9	Sequence 9, Appl
32	80.5	6.4	453	6	US-10-453-372-106	Sequence 106, Appl
33	80.5	6.4	453	6	US-11-154-673-9	Sequence 9, Appl
34	80.5	6.4	550	7	US-11-154-673-7	Sequence 8, Appl
35	80.5	6.4	625	7	US-11-154-673-7	Sequence 7, Appl
36	79.5	6.3	343	7	US-11-087-099-7967	Sequence 7967, Ap
37	79	6.3	164	7	US-11-031-206-84	Sequence 84, Appl
38	78.5	6.3	161	7	US-11-031-206-96	Sequence 96, Appl
39	77	6.1	206	6	US-10-498-026-80	Sequence 80, Appl
40	76.5	6.1	164	7	US-11-031-206-82	Sequence 82, Appl
41	76.5	6.1	253	6	US-10-467-657-55684	Sequence 5684, Ap
42	76	6.1	453	6	US-10-878-556A-141	Sequence 141, App
43	75.5	6.0	373	7	US-11-087-099-1660	Sequence 1660, App
44	75.5	6.0	373	7	US-11-087-099-5200	Sequence 5200, Ap
45	75.5	6.0	486	6	US-10-467-962B-59	Sequence 59, Appl

## ALIGNMENTS

RESULT 1	US-10-527-771-10		
Sequence 10, Application US/10527771			
Publication No. US20050271683A1			
GENERAL INFORMATION:			
APPLICANT: University Gent			
TITLE OF INVENTION: Ostertagia vaccine			
FILE REFERENCE: 2002-015			
CURRENT FILING DATE: 2005-03-11			
PRIOR APPLICATION NUMBER: US 10/243,319			
PRIOR FILING DATE: 2002-09-13			
NUMBER OF SEQ ID NOS: 27			
SOFTWARE: PatentIn version 3.2			
SEQ ID NO 10			
LENGTH: 231			
TYPE: PRT			
ORGANISM: Ostertagia ostertagi			
US-10-527-771-10			
Query Match	100.0%	Score 1253;	DB 6; Length 231;
Best Local Similarity	100.0%	Pred. No. 3.8e-123;	
Matches 231; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSAAVVAVVLTALFSYAENGFCPCPSLSQSDARQIFLDFHNDVRNTALGGLINMTN	60
DB	1	MSAAVVAVVLTALFSYAENGFCPCPSLSQSDARQIFLDFHNDVRNTALGGLINMTN	60
QY	61	ADAVILGPANMYKYDWDONLEEVAAQOIAPCNDLPIINTSLAONIAWLYFKDSEETV	120
DB	61	ADAVILGPANMYKYDWDONLEEVAAQOIAPCNDLPIINTSLAONIAWLYFKDSEETV	120
QY	121	LOQVSWYVVSASLGFPMKGTGLDQFANQWAEPLANTANTRKRVGCAHKICPAQNNVVC	180
DB	121	LOQVSWYVVSASLGFPMKGTGLDQFANQWAEPLANTANTRKRVGCAHKICPAQNNVVC	180
QY	181	VYGSFKLAPNEVIMQGRACVCDARPDSCDNLCDTRDAAVRRHCCASP	231
DB	181	VYGSFKLAPNEVIMQGRACVCDARPDSCDNLCDTRDAAVRRHCCASP	231
RESULT 2	US-10-527-771-2		
Sequence 2, Application US/10527771			
Publication No. US20050271683A1			
GENERAL INFORMATION:			
APPLICANT: University Gent			
TITLE OF INVENTION: Ostertagia vaccine			

```

? FILE REFERENCE: 2002-015
? CURRENT APPLICATION NUMBER: US/10/527,771
? CURRENT FILING DATE: 2005-03-11
? PRIOR APPLICATION NUMBER: US 10/243,319
? PRIOR FILING DATE: 2002-09-13
? NUMBER OF SEQ ID NOS: 27
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 2
? LENGTH: 236
? TYPE: PRT
? ORGANISM: Ostertagia ostertagi
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (37)..(37)
? OTHER INFORMATION: The xaa at location 37 stands for Lye, or Gln.
? FEATURE:
? NAME/KEY: misc feature
? LOCATION: (121)..(121)
? OTHER INFORMATION: The xaa at location 121 stands for Asn, or Thr.
? US-10-527-771-2

```

FILE REFERENCE: 110.01860101  
CURRENT APPLICATION NUMBER: US/10/515,868  
CURRENT FILING DATE: 2004-11-24  
PRIOR APPLICATION NUMBER: 60/383,628  
PRIOR FILING DATE: 2002-05-28  
PRIOR APPLICATION NUMBER: PCT/US03/16669  
PRIOR FILING DATE: 2003-05-28  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 5  
LENGTH: 243  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-10-515-868-5

Query Match  
Best Local Similarity 21.6%; Score 119; DB 6; Length 243;  
Matches 42; Conservative 29; Mismatches 85; Indels 38; Gaps 8;

Qy 41 HNDVRNIALGNGLIINTVNADAVILGPAQNMVYKVDMDNLEEVAAQIAPC-----ND 94  
Db 45 HNELRQVS-----PQSNILIKEMVYQAANNAQKANNCLLHSSST 87  
Qy 95 PLPITSLAQNIARMLYFKDSEETVLQGVSWYVWSASIGPKGTLDQFANQMAEPLAN 154  
Db 88 DRKINIKGCEML--YMSDPTSMRTVIQ--SWYBENENFVGVGAKPNSAVGHYQ---- 139  
Qy 155 IANYNRKVGCAHKICPAQNMNV-----SCYVSGPKAPNEVINOEGKACV-CDARPS 208  
Db 140 LWMYSSPKGCGVAVCPMDTLKPYVCHYCPGWN-NVWKSKTPYHQTPCASCPCNNCDN 198  
Qy 209 FCCDMLCDTRDAAS 222  
Db 199 GLCTNSCDPEDLS 212

RESULT 6  
US-11-153-222A-3  
Sequence 3, Application US/11153222A  
Publication No. US20060039921A1  
GENERAL INFORMATION:  
APPLICANT: Lustigman, Sara  
TITLE OF INVENTION: Adjuvancy and Immune Potentiating Properties of Natural Products  
TITLE OF INVENTION: of Onchocerca Volvulus  
FILE REFERENCE: 454-37  
CURRENT APPLICATION NUMBER: US/11/153,222A  
CURRENT FILING DATE: 2005-06-15  
PRIOR APPLICATION NUMBER: 60/580,254  
PRIOR FILING DATE: 2004-06-15  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 3  
LENGTH: 220  
TYPE: PRT  
ORGANISM: Onchocerca volvulus  
US-11-153-222A-3

Query Match  
Best Local Similarity 9.4%; Score 117.5; DB 7; Length 220;  
Matches 52; Conservative 27; Mismatches 117; Indels 29; Gaps 9;

Qy 4 AVVAVVLLALPSYALAGFCPCPSLSQSDSARQIFLDFHNDVRNIALGNGLIINTVNADA 63  
Db 9 AIIYAV-----TGHDCHRG-KLTSLQORDIIDEHNKRSRLVKGN-----FANKDG 53  
Qy 64 VILGPAQNMVYKVDMDNLEEVAAQIAPC-----NDPLPITSLAQNIARMLYFKDSE--E 117  
Db 54 NSMPGKXKMEMENHCELEISQNMADOCIFGYSBENREGGENTYALGLPKDVEVFT 113  
Qy 118 ETVLQGVSWYVWSASIGPK--GTLD-QFANQMAEPLANIANYNRRKVGCAHKI-CPAQ 173  
Db 114 SAALFAISWMTELIRSYRNPNKLTSSVASQDVLHFTQAMWGTHKVGCGIANGCDG 173

Qy 174 QNMVVSQVYSGPKLAPNEVINOEGKACVCDARPSFCCDMLCDTR 218  
Db 174 EAFIVCHYAPRAGNTIGELIYQSGSPC-----KVNHCCKTKCSRK 214

RESULT 7  
US-11-090-439-38  
Sequence 38, Application US/11090439  
Publication No. US20050266442A1  
GENERAL INFORMATION:  
APPLICANT: Squillace, Rachel  
APPLICANT: Weiner, Michael P.  
TITLE OF INVENTION: Immortalized Human Tuberosus Sclerosis Null  
TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof  
FILE REFERENCE: 24318-502  
CURRENT APPLICATION NUMBER: US/11/090,439  
CURRENT FILING DATE: 2005-03-25  
PRIOR APPLICATION NUMBER: 60/556,344  
PRIOR FILING DATE: 2004-03-25  
NUMBER OF SEQ ID NOS: 62  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 38  
LENGTH: 258  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-11-090-439-38

Query Match  
Best Local Similarity 8.9%; Score 111.5; DB 7; Length 258;  
Matches 53; Conservative 22; Mismatches 62; Indels 99; Gaps 12;

Qy 27 LSQSDSARQIFLDFHNDVRNIALGNGLIINTVNADAVILGPAQNMVYKVDMDNLEEVAA 86  
Db 62 ISQNDMT--AIIIDYNOVR-----GVPPPAAMETVWMDENLAKSNE 102  
Qy 87 QOIAPC--NDPLPITSLAQNIARMLYFKDSEETVLQGVSWYVWSASIGPKGTLD 142  
Db 103 AMAATCINDHGSYSLARLQNLGNS--VRGRYRSIIQLYKPY-----D 144  
Qy 143 QFANQMAEPLANIANYR-----NRKVGCAHKICPAQNMVVSQVY 182  
Db 145 E-VKQYAPFPQDCNCPRCFPMCHYTMWATSNRIGCAIHTC---QNM--NVW 197  
Qy 183 GSPKLAPEVINOEGKACVCDARP-----DSFCCDMLC 215  
Db 198 GS-----VWRAVYLVCNYPAPKGNWIGAPYKVGVPCCSPSYGSGCTDNLG 245

RESULT 8  
US-10-515-868-6  
Sequence 6, Application US/10515868  
Publication No. US20050282729A1  
GENERAL INFORMATION:  
APPLICANT: Hamilton, David W  
APPLICANT: Roberts, Kenneth P  
APPLICANT: Ensrud, Kathy M  
TITLE OF INVENTION: CRISP POLYPEPTIDES AS CONTRACEPTIVES AND INHIBITORS OF SPERM  
TITLE OF INVENTION: CAPACITATION  
FILE REFERENCE: 110.01860101  
CURRENT APPLICATION NUMBER: US/10/515,868  
CURRENT FILING DATE: 2004-11-24  
PRIOR APPLICATION NUMBER: 60/383,628  
PRIOR FILING DATE: 2002-05-28  
PRIOR APPLICATION NUMBER: PCT/US03/16669  
PRIOR FILING DATE: 2003-05-28  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 6  
LENGTH: 245  
TYPE: PRT  
ORGANISM: HOMOSAPIEN  
US-10-515-868-6

	Query Match	8.5%	Score 106.5;	DB 6;	Length 245;
	Best Local Similarity	23.6%	Pred. No. 0.0011;		
	Matches	54;	Conservative	25;	Mismatches 99; Indels 51; Gaps 10;
QY	28	SQSDSARQIFLDFHNDVRNATLNGLLINMYNADAVILGPRQNMYKYDMDNCNLEEVAAQ	:	:	:
		: : : :	:	:	:
Db	34	TOTOVOREI-VNKHELRRAVS-----PPRRNLTKEMNEEAANAOK	:	:	:
		: : : :	:	:	:
QY	88	QIAPCN----DLPINTSL--AQNIARWLFEKDSSEETVLQVSYWYSASLGFMKGITL	:	:	:
		: : : :	:	:	:
Db	76	WANCTNYHSHPKDMSTLKGEN-----LYMSASSMSQAISWFDEYNDDFVGIPRT	:	:	:
		: : : :	:	:	:
QY	142	DQFANQMAEPLANTIANRYNRKKYGCHHKCPRAQ--NMVYSCYGSFKLAPR- -VIWQEG	:	:	:
		: : : :	:	:	:
Db	132	--PAVVGHYTQVWVSSYLWGCGNAACPNQIKLTYYVCQYCPAGMANKLLIVPYEGG	:	:	:
		: : : :	:	:	:
QY	198	KACV--CDARPDSFCDDNLCDTRDAAS-----VRHOCCAS	:	:	:
		: : : :	:	:	:
Db	189	APCASCPNDCDGLCTNGCKGYEDLSNCKSLTLTLCKHQLRVDSCKAS	:	:	:
		: : : :	:	:	:

```

RESULT 9
US-11-067-573-2
; Sequence 2, Application US/11067573
; Publication No. US20050260622A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin
; APPLICANT: Godowski, Paul
; APPLICANT: Wood, William
; APPLICANT: Smith, Victoria
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES
; TITLE OF INVENTION: AND NUCLEIC ACIDS ENCODING THE SAME
; FILE REFERENCE: 39780-343ORIC34C1
; CURRENT APPLICATION NUMBER: US/11/067,573
; PRIORITY FILING DATE: 2005-02-25
; PRIOR APPLICATION NUMBER: US 10/198,766
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: US 10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/380,138
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: PCT/US99/05028
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: US 60/087,098
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-573-2

Query Match      8.3%; Score 103.5; DB 7; Length 500;
Best Local Similarity    22.2%, Pred. No. 0.0033;
Matches    50; Conservative     26; Mismatches   68; Indels    81; Gaps    13;

DQ          30 SDSARQIFDPFNDVNRNATLNGELINWTVADAVILGPAONMYKVDMDCNIEEVAQAQI 89
           :|::|||::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
DB 58 TINDMQSIIDLNKKR-----SQYYPSASMEYWTVDVELERSAESWA 100

DY          90 APC---NDPLPNTSLAQNI-ARMLYFKDSSEETYLQOVSWTWVASLSGFNMGTLDLPFA 145
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:
DB 101 ESLTMEHGASALPSIGONIAGHWGRYR----PPTHVSQMW-----DEVKDF 145

DY          146 NWGAEP LANIANVR-----NNRVGAHHICPAQQM-----VV 178
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:
DY          146 YPEHEHCANPCPRCGSPCTHTYTQVWATSNRGAINLC---HNMINIQGIWKAVYL 202
           :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~:
```

```

Qy      179 SCYLGSPK-----LAPNEVIWQEGKACVCDARPSF---CCDNLC 215
          |||      |||      |||      |||      |||      |||
Db      203 VCNV-SPKGNWGHAP---YKHGRP--CSACPSPFGGCGRENTLC 240

```

```

RESULT 10
US-10-515-866-4
; Sequence 4, Application US/10515868
; Publication No. US20050282729A1
; GENERAL INFORMATION:
; APPLICANT: Hamilton, David W
; APPLICANT: Roberts, Kenneth P
; APPLICANT: Ensrud, Kathy M
; TITLE OF INVENTION: CRISP POLYPEPTIDES AS CONTRACEPTIVES AND INHIBITORS OF SPERM
; TITLE OF INVENTION: CAPACITATION
; FILE REFERENCE: 110.01860101
; CURRENT APPLICATION NUMBER: US/10/515,868
; CURRENT FILING DATE: 2004-11-24
; PRIOR APPLICATION NUMBER: 60/183,628
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: PCT/US03/16669
; PRIOR FILING DATE: 2003-05-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 243
; TYPE: prt
; ORGANISM: HOMOSAPIEN
US-10-515-866-4

```

	Query Match	7.9%	Score 69;	DB 6;	Length 243;		
	Best Local Similarity	22.8%	Pred. NO.	0.0065;			
	Matches	47;	Conservative	26;	Mismatches	97;	
				Indele	36;	Gaps	10.
Qy	28	SQSDSARQIFLDFNDVRRNTALNGLINMTVNADAVILGPQNMYKYDWDNCNLEBVAQAQ	87				
Db	32	TOLQVOREI-VYKNEHLRKAAS-----PPASNNLLKNESREVTYNQR	73				
Qy	88	GIAPC----NDPLPINTS--LAONIAWLYFPDSSEETVLQOVSYWYSASIGFPMKTGL	141				
Db	74	WANCKTLGHSDDEBKRTITRCGEN---LYM--SSDPFSWSAISOXWDELIDFLFYGVG-	126				
Qy	142	DQFANQMAEPLIANINAYNRKKVGCAHKICPAQOQM--VVSCTY--GSFKLANEVIMQEG	197				
Db	127	PKSPNAVVGHYTQLTWYSTGYQGCGCIACYPNDISLKYYTVYCQCPAGNNMRKNTPTPOOG	186				
Qy	198	KACY-CDAAPDSFCDDNLCDTRDAAS	222				
Db	187	TPCAGCPDDCCDKGLCTNSCYQYDLIS	212				

```

RESULT 11
US-10-506-443A-41
; Sequence 41, Application US/10506443A
; Publication No. US20060013817A1
; GENERAL INFORMATION:
; APPLICANT: Sahin Dr., Ugur
; APPLICANT: Tureci Dr., Ozlem
; APPLICANT: Koslowski Dr., Michael
; TITLE OF INVENTION: Genetic Products Differentially Expressed in Tumors and Use There
; FILE REFERENCE: 342-3PCT
; CURRENT APPLICATION NUMBER: US/10/506,443A
; CURRENT FILING DATE: 2004-09-01
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-506-443A-41

Query Match          7.9%; Score 99; DB 6; Length 243;
Best Local Similarity 22.8%; Pred. No. 0.0065;

```

Matches	47;	Conservative	26;	Mismatches	97;	Indels	36;	Gaps	10;
QY	28	SQSDBARQIFLDHNDVNRNTALNGLNWTYNADAVILGPQNNYKYDWDNCELEVAQAQ							87
Db	32	TOLQVOREI-VKIKHELRKVAS-----PPASNNLKNWMSREVTTNAQR							73
QY	88	QIAPC-----NDPLPINTS--LAQNIARWLTFDSEETVLQOVSWYVWSASIGFMKGTKL							141
Db	74	WANKTCLQHSBEDKDKTSTRGSEN----LYM--SSDPLTSMSSALQSWDELIDLPFYGGV							126
QY	142	DQFANQMAEPLANTIANYNRKYGCNHHKTCPAQNM--VVSCTY--GSEPLAPNEVYWGEG							197
Db	127	PKSPAAVVGHYQTLVWYSTQYQGCIAIACPNDSLKTYVVCQCPAGNNMRKMTPYQGG							186
QY	198	KACV--CDARPDSPCCDNLCDTRDAAS							222
Db	187	TPCAGCPRDDCKGLCTNSCQYQDPLS							212

```

RESULT 12
US-10-527-500-23
; Sequence 23, Application US/10527500
; Publication No. US20060004186A1
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
; APPLICANT: REPRESENTED BY THE SECRETARY OF HEALTH AND
; APPLICANT: HUMAN SERVICES
; APPLICANT: Valenzuela, Jesus G.
; APPLICANT: Ribetiro, Jose M.C.
; APPLICANT: Kamawel, Shaden
; APPLICANT: Belkaid, Yasmine
; APPLICANT: Fischer, Laurent Bernard
; APPLICANT: Audoumet, Jean-Christophe
; APPLICANT: Milward, Francis William
; TITLE OF INVENTION: P. ARIASI POLYPEPTIDES AND P. PERRICIOSUS POLYPEPTIDES AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 4239-66903-02
; CURRENT APPLICATION NUMBER: US/10/527,500
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: PCT/US2003/029833
; PRIOR FILING DATE: 2003-09-18
; PRIOR APPLICATION NUMBER: US 60/425,852
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/412,327
; PRIOR FILING DATE: 2002-09-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 23
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Phlebotomus arlasi
US-10-527-500-23

```

[illegible]

```

QY      196  EGKAC-VCDARPDFCCDNLCDTRD 211
          ||| | |
Db      234  TGKPCSQCQGGKKCDSDVYKNLCDASE 256

```

```

RESULT 13
US-10-453-372-100
; Sequence 100, Application US/10453372
; Publication No. US20060003323A1
; GENERAL INFORMATION:

```

```

: TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
: FILE REFERENCE: 21402-589 A
: CURRENT APPLICATION NUMBER: US/10/453,372
: CURRENT FILING DATE: 2003-06-03
: PRIOR APPLICATION NUMBER: 09/789390
: PRIOR FILING DATE: 2001-02-23
: PRIOR APPLICATION NUMBER: 60/185967
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: 09/823187
: PRIOR FILING DATE: 2001-03-29
: PRIOR APPLICATION NUMBER: 60/195792
: PRIOR FILING DATE: 2000-03-10
: PRIOR APPLICATION NUMBER: 09/839446
: PRIOR FILING DATE: 2001-03-19
: PRIOR APPLICATION NUMBER: 60/199476
: PRIOR FILING DATE: 2000-03-25
: PRIOR APPLICATION NUMBER: 09/863776
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: 60/208263
: PRIOR FILING DATE: 2000-05-31
: PRIOR APPLICATION NUMBER: 09/939398
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: 60/227800
: PRIOR FILING DATE: 2000-08-25
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 1609
: SOFTWARE: Curaseq1st version 0.1
: SEQ ID NO 100
: LENGTH: 497
: TYPE: PRT
: ORGANISM: Homo sapiens
: OS-10-453-372-100

```

Query Match	7.4%	Score 92.5;	DB 6;	Length 497;
Best Local Similarity	21.0%	Pred. No. 0.074;		
Matches 50;	Conservative 28;	Mismatches 91;	Indels 69;	Gaps 12

```

QY 10 LIALFEYARAGCCPSLSDSAQIFLDFNDVNRRIALGNLMTVAADVILGPA 69
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 LLSKYOHNSHRVRRIAPRED-KEELIIMLHNKLRGOV-----PQA 74
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 70 OMNYKYVDMDCNLEEVAAQOIAFC--NDPLEPINTSLAONI-ARWLTKDESEBTVLQOVS 125P
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 75 SNNEVYTWMDDELEKSAAMAAWASQCIEMHEGPTSLVLSIGQNLGAHNGRRYRSPQ----FHVQS 130Q
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 126 WY-----WVSASLGFPMKGYKLDOEPANOMAEPLIANIYTRNRKYGCHAKI 168P
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 131 WYDEVVDYTYPYPSBGNPCPERCSGPMCTHYTOIV-WA-----TTNKIGCAVNT 179P
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 170 CPAQOMNVY-----SCYVSGPKLAP-----NEVIMQEGKAC-VCDARPDSPFCDDNIC 215
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 C--RRQTVGWEIENAAVYFVCCNPKYKGNMIGEARPKYKGRCSBSPSYGSCCANNNIC 234
   | : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 14
US-10-453-372-102
; Sequence 102, Application US/10453372
; Publication No. US2006000323A1
; GENERAL INFORMATION:
; APPLICANT: Alsbjork, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES,
; FILE REFERENCE: 21402-589 A
; NUCLEIC ACIDS ENCODING SAME, AND METHOD

```

CURRENT APPLICATION NUMBER: US/10/453,372  
CURRENT FILING DATE: 2003-06-03  
PRIOR APPLICATION NUMBER: 09/789390  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/185967  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/823187  
PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195792  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476  
PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
PRIOR FILING DATE: 2000-08-25  
Remaining Prior Application data removed - See file Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: Curaseqlist version 0.1  
SEQ ID NO 102  
LENGTH: 497  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-453-372-102

Query Match 7.4% Score 92.5; DB 6; Length 497;  
Best Local Similarity 21.0%, Pred. No. 0.074; Indels 63; Gaps 12;  
Matches 50; Conservative 28; Mismatches 91; Indels 63; Gaps 12;  
QY 10 LIALPSYAEAGFCPCNSLSQSDSARQIFLDPHNDVRRNIALGNGLIMTVADAVILGPA 69  
DB 34 LLSKYQHNSHSHRVRAIRPRED-KEITLMLHNRKRGQVQ-----PQA 74  
QY 70 QNMKYVDWDCNLEEVAAQIAPC--NDPLPINTSLAONT-ARWLFPKDSSEETVLQOVS 125  
DB 75 SNMEYMTWDBLEKSAAMAASQCIWEHGPTSLVLSIGNLGAHWGRYSPG---FHVQS 130  
QY 126 WY-----WVSASLGFMKGTCLDQFANQWAEPLANIANRYNRKYGCAHKI 169  
DB 131 WYDEVKDYTPYPBSCNFWCPERCSCGPMCTHTYQIV-WA-----TTNKIGCAVNT 179  
QY 170 CPAQNMVY-----SCVYGSPLAP-----NEVIWQBGKAC-VCDARPDSCDNLG 215  
DB 180 C---RKMTVWGEVWENAVYFVCNYSFKGNWIGEAPYKNGRPSCECPSPYSGSCRNNLC 234

RESULT 15  
US-10-453-372-98  
Sequence 98, Application US/10453372  
Publication No. US20060003323A1  
GENERAL INFORMATION:  
APPLICANT: Alcobrook, et al.  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-589 A  
CURRENT APPLICATION NUMBER: US/10/453,372  
CURRENT FILING DATE: 2003-06-03  
PRIOR APPLICATION NUMBER: 09/789390  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/185967  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: 09/823187  
PRIOR FILING DATE: 2001-03-29  
PRIOR APPLICATION NUMBER: 60/195792  
PRIOR FILING DATE: 2000-03-10  
PRIOR APPLICATION NUMBER: 09/839446  
PRIOR FILING DATE: 2001-03-19  
PRIOR APPLICATION NUMBER: 60/199476

PRIOR FILING DATE: 2000-03-25  
PRIOR APPLICATION NUMBER: 09/863776  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: 60/208263  
PRIOR FILING DATE: 2000-05-31  
PRIOR APPLICATION NUMBER: 09/939398  
PRIOR FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 60/227800  
PRIOR FILING DATE: 2000-08-25  
Remaining Prior Application data removed - See file Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1609  
SOFTWARE: Curaseqlist version 0.1  
SEQ ID NO 98  
LENGTH: 475  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-453-372-98

Query Match 7.3% Score 91.5; DB 6; Length 475;  
Best Local Similarity 21.0%, Pred. No. 0.089; Indels 69; Gaps 12;  
Matches 50; Conservative 28; Mismatches 91; Indels 69; Gaps 12;  
QY 10 LIALPSYAEAGFCPCNSLSQSDSARQIFLDPHNDVRRNIALGNGLIMTVADAVILGPA 69  
DB 12 LLSKYQHNSHSHRVRAIRPRED-KEITLMLHNRKRGQVQ-----PQA 52  
QY 70 QNMKYVDWDCNLEEVAAQIAPC--NDPLPINTSLAONT-ARWLFPKDSSEETVLQOVS 125  
DB 53 SNMEYMTWDBLEKSAAMAASQCIWEHGPTSLVLSIGNLGAHWGRYSPG---FHVQS 108  
QY 126 WY-----WVSASLGFMKGTCLDQFANQWAEPLANIANRYNRKYGCAHKI 169  
DB 109 WYDEVKDYTPYPBSCNFWCPERCSCGPMCTHTYQIV-WA-----TTNKIGCAVNT 157  
QY 170 CPAQNMVY-----SCVYGSPLAP-----NEVIWQBGKAC-VCDARPDSCDNLG 215  
DB 158 C---RKMTVWGEVWENAVYFVCNYSFKGNWIGEAPYKNGRPSCECPSPYSGSCRNNLC 212

Search completed: March 3, 2006, 19:14:24  
Job time : 22 secs





R:Stulston, J.  
submitted to the EMBL Data Library, November 1995  
A:Accession: T24494  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-221 <MIL>  
A:Cross-references: UNIPROT:P90559, UNIPARC:UPI000008266B, EMBL:Z68108, PIDD:CAA92137.1  
A:Experimental source: clone T05A10  
C:Genetics:  
A:Gene: CESP:T05A10.4  
A:Map position: X  
A:introns: 18/3; 54/1; 106/3; 142/2; 190/3

Query Match 14.2%; Score 178.5; DB 2; Length 221;  
Best Local Similarity 29.3%; Pred. No. 2.4e-08;  
Matches 51; Conservative 35; Mismatches 59; Indels 29; Gaps 8;

QY 22 CCRNLSQSDSARQIFLDFHNDVRNIALG-----NGLINMTVNADAVILGPAQNMVYKD 76  
DB 61 CNKSTITQLQ--QETILTHNELRSLAPGKQNRKRLMN-----GARNYKLD 107  
QY 77 WDCNLEEVAAQOIAFCND---PLPINTSLAQNIAFWLYFKDSEERTV-LQOVSWYWSAS 132  
DB 108 WDCELASLAAWSTSCPGHFMPQSVLGSVAQLFKRFYFPGDHDSTVHMRNAMYMQO- 166  
QY 133 LGFMKGTLDQ---FANQMAEPLNANTYRNKRYGCAHKICPAQNMVYSCY 182  
DB 167 -GEEKGNE-DQKNRFYARNYFGMANMAKGYRVCGYMCGDESLFTCLY 218

## RESULT 3

T27833  
hypotheetical protein ZK384.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
R:Almscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20427  
A:Accession: T27833  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-196 <MIL>  
A:Cross-references: UNIPROT:O62507, UNIPARC:UPI000007B62C, EMBL:Z82092; PIDD:CA805010.1  
A:Experimental source: clone ZK384  
C:Genetics:  
A:Gene: CESP:ZK384.1  
A:Map position: 5  
A:introns: 58/3; 95/3; 157/2  
C:Superfamily: yellowjacket venom allergen antigen 5

Query Match 13.3%; Score 167; DB 2; Length 196;  
Best Local Similarity 26.1%; Pred. No. 2.1e-07;  
Matches 63; Conservative 30; Mismatches 70; Indels 78; Gaps 12;

QY 5 VVVAVLALFSYAAGFCCPNLSQSDSARQIFLDFHNDVRNIALGNGLINMTVNADAV 64  
DB 3 LILFLLLAITS--SSGQLSPN-----GRQVLDLDFHNKLRSGVALG-----VFSANGT 47  
QY 65 ILGPAQNMVYKVDNLEEVAAQOIAFCNDPL--PINTSLAQNIAFWLYFKDSEERTV 122  
DB 48 IKPPARNMERLTGQGFERLAQDYVADCPDGLPIGRNIGM----- 90  
QY 123 QVSWYWSASLGFMKGTLDQFANQMAE-----LANIANYN-----RKYG 164  
DB 91 ----YT-----TKVIDALNDMAEFOVNGMLSTIYNDISIGAAAGWAGTKYVG 137  
QY 165 CAHKICPAQNMVYSCYVSGPKLAPNEVIWEGKAC-----VCDARPDSFCCD---NL 214  
DB 138 CGYKRCQD-PINVVVYVCGYVQGNLVRPIYKEGPPCTACPPMRICPGQKE--CCDRVMGL 194  
QY 215 C 215

DB 195 C 195

## RESULT 4

T33397  
hypotheetical protein B0545.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T33397  
R:Rolling, T.  
submitted to the EMBL Data Library, July 1998  
A:Description: The sequence of C. elegans cosmid B0545.  
A:Reference number: Z21337  
A:Accession: T33397  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-241 <ROH>  
A:Cross-references: UNIPROT:O76663, UNIPARC:UPI000007723F, EMBL:AF078781; PIDD:AAC26915.  
A:Experimental source: strain Bristol N2; clone B0545  
C:Genetics:  
A:Gene: CESP:B0545.3  
A:Map position: 4  
A:introns: 38/3; 68/3; 123/3; 155/3; 194/2

Query Match 12.8%; Score 160; DB 2; Length 241;  
Best Local Similarity 29.0%; Pred. No. 1.1e-06;  
Matches 47; Conservative 27; Mismatches 50; Indels 38; Gaps 8;

QY 34 ROIPLDFHNDVRNIALGNGLINMTVNADAVILGPAQNMVYKVDNLEEVAAQOIAFCN 93  
DB 33 ONLILDKNEIRSGVALG---QYAVDD--YLPPADMYGLDWDCLELBAQOQPAQCN 86  
QY 94 DPLPINTSLAQNIAFW-----LYFKDSE-----EETVLQOVSWYWSASLGFMKGTLD 141  
DB 87 --LQKENS-GRQMGWDEVRGENAFYRTTGLDVSAGVALGICQMGEBIAMIAGIKNKL 143  
QY 142 DQFANQ-----WAEPLNANTYRNKRYGCAHKICPAQNM 175  
DB 144 SRYDSRIGHAQLIMKE-----TKKGCAGVQECPARQD 176

## RESULT 5

T27834  
hypotheetical protein ZK384.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
R:Almscough, R.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z20427  
A:Accession: T27834  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-212 <MIL>  
A:Cross-references: UNIPROT:O62508, UNIPARC:UPI000007591B, EMBL:Z82092; PIDD:CA805011.1  
A:Experimental source: clone ZK384  
C:Genetics:  
A:Gene: CESP:ZK384.2  
A:Map position: 5  
A:introns: 60/3; 106/3; 180/2  
C:Superfamily: yellowjacket venom allergen antigen 5

Query Match 11.7%; Score 146; DB 2; Length 212;  
Best Local Similarity 24.7%; Pred. No. 1.6e-05;  
Matches 56; Conservative 28; Mismatches 81; Indels 62; Gaps 9;

QY 24 PNLISQ---SDSARQIFLDFHNDVRNIALGNGLINMTVNADAVILGPAQNMVYKVDN 80  
DB 12 PNLICDFRGRGPTRAQREIVDFHNSIRSLANGYV-----DGVKPPAKDMKMKMDPI 65  
QY 81 LEEVAAQOIAFCNDPLPINTSLAQN-----IARMLYFKDSEERTVILQOVSWY 128

Db 66 LAGMANNAATCPSTLTDSKMLGRNYHRLANVTSGSLDKVALFAVKMERQFOERGW-- 123  
 QY 129 VSASLGFMKGTLDOPANO-----WAEPLANIANYRKRKGCARHKICPAQOM-- 176  
 Db 124 -----KNGEPRFEGDRRLTSATQYWA-----TRHVGCGVNCIDAEKNLPG 166  
 QY 177 -----VSCVYSPKLAPEVIMQEGKACVCDARPDSPCCD--NLG 215  
 Db 167 YANKVAVVICEYQSKGNIHGLPIYKSGP--TCSACPSATGTCERRSGLC 211

## RESULT 6

T22438

hypoetical protein F49B11.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T22438

R:Baynes, C.

A:Submitted to the EMBL Data Library, March 1996

A:Reference number: Z19564

A/Accession: T22438

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-207 &lt;MIL&gt;

A:Cross-references: UNIPROT:Q93746; UNIPARC:UPI0000077498; EMBL:Z70308; PDB:CAA94350.1;

C:Experimental source: clone F49B11

C:Genetics:

A:Gene: CESP:F49B11.6

A:Map position: 4

A:Introns: 58/3; 104/3; 176/2

C:Superfamily: yellowjacket venom allergen antigen 5

Query Match 11.4%; Score 142.5; DB 2; Length 207;  
 Best Local Similarity 23.6%; Pred. No. 3.2e-05;  
 Matches 56; Conservative 19; Mismatches 79; Indels 83; Gaps 12;  
 QY 22 CCPNLSQ-SDSARQIFLDFHNDVRNIALGNGLIWTVNADAVILGPAQNNYKYVMDGN 80  
 Db 10 CLAGVFSQPTSTGQAIYDAHAKKSSIAKGTYYAKTQKS-----GSNRKTKMDAT 63  
 QY 81 LEEVAQAQIAPCNDPLPINTSLAQNIAKRLYFKDSEETVLQOVSNVYVWASLGFMKGT 140  
 Db 64 VATAQNVANVC-----PTGHSQSGVGENLY-----WYWTSGTIG-----N 100  
 QY 141 LDPF-----ANQ-WAEPLANIANYRKRKGC 165  
 Db 101 LDTFPAASSSWSEFQGYGWTSTNTLDMNTFNTGIGHATQMAWMTFA-----IGC 151  
 QY 166 AHKIC---PAQ--QNNVYSCVYSPKLAPEVIMQEGKACVCDARPDSPCCD--LC 215  
 Db 152 GVNKGCKDPSNGYNKVAVVCYKTPGNLYNPPIYQGGTTCA--ACPSGTACDSGLC 206

## RESULT 7

T22439

hypoetical protein F49B11.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T22439

R:Baynes, C.

A:Submitted to the EMBL Data Library, March 1996

A:Reference number: Z19564

A/Accession: T22439

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-213 &lt;MIL&gt;

A:Cross-references: UNIPROT:Q93747; UNIPARC:UPI000007849D; EMBL:Z70308; PDB:CAA94351.1;

C:Experimental source: clone F49B11

C:Genetics:

A:Gene: CESP:F49B11.4

A:Map position: 4

A:Introns: 14/1; 60/3; 105/3; 175/2

C:Superfamily: yellowjacket venom allergen antigen 5

Query Match 10.8%; Score 135.5; DB 2; Length 213;  
 Best Local Similarity 25.2%; Pred. No. 0.00014;  
 Matches 59; Conservative 26; Mismatches 92; Indels 57; Gaps 12;

QY 9 VLLALPSYAEAGFCCPNLSQSDSARQIFLDFHNDVRNIALGNGLIWTVNADAVILGP 68  
 Db 4 VLLALPSA-IGVMSDNF-----SKEGQNLINLVHNEFRSQLAQG-----LSFRGVKPS 53  
 QY 69 AQNTKYVMDCNLEEVAAQAIAPCNDPLPINTSLAQNIAKRLYFKDS-----EETVQ 122  
 Db 54 ASMRKISWSKKLTJNAATKFAETC-----PKNHSVVMNTGESIFMFISSSLSTPEQYATLA 109  
 QY 123 QVSWYVWASLGFMKGTLDOP-ANQWAEPLANIANYR-----NRKYGCARH 168  
 Db 110 PQKMW-----NEETNGMDSLIYNHNSQRFQIGHANQVAMHTTSKVGCGS 155  
 QY 169 ICPA---QNNVYSCVYSPKLAPEVIMQEGKACV-CDARPDSP--CCDNLCD 216  
 Db 156 KCAVGTPEQTMVVCRYFQKNIIEGEPYINBETCTKC--PEBYQKCPSGLC 206

## RESULT 8

T22432

hypoetical protein F49B11.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T22432

R:Baynes, C.

A:Submitted to the EMBL Data Library, March 1996

A:Reference number: Z19564

A/Accession: T22432

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-207 &lt;MIL&gt;

A:Cross-references: UNIPROT:Q20603; UNIPARC:UPI000007D17A; EMBL:Z70308; PDB:CAA94344.1;

C:Experimental source: clone F49B11

C:Genetics:

A:Gene: CESP:F49B11.10

A:Map position: 4

A:Introns: 58/3; 104/3; 175/2

C:Superfamily: yellowjacket venom allergen antigen 5

Query Match 9.7%; Score 121; DB 2; Length 207;  
 Best Local Similarity 23.0%; Pred. No. 0.0025;  
 Matches 56; Conservative 25; Mismatches 89; Indels 74; Gaps 10;

QY 5 VVAVVTLALPSYAEAGFCCPNLSQSDSARQIFLDFHNDVRNIALGNGLIWTVNADAV 64  
 Db 4 LILVVALAVGSAADPG-----SSGNGCIINAHTTSLAKG-----TVAAKT 47  
 QY 65 ILGPAQNNYKYVMDCNLEEVAAQAIAPCNDPLPINTSLAQNIAKRLYFKDSEETVLQOV 124  
 Db 48 QKSPGTNLLKKMWD---SAVAAS-----AQVYANCCPTGHSADAGLGENL 89  
 QY 125 SWYVWASLGFMKGTLDOPANO-----WAEPLANI-----ANY 158  
 Db 90 YWYWTSGSLG-----DLNQGSASASWEKEFQDYGKSNMTLIDLFTGTGIGHATQYWA 144  
 QY 159 RNRKYGCARHKICPAQON-----WVYSCVYSPKLAPEVIMQEGKACVCDARPDSPCCD-- 212  
 Db 145 KSNLIGCGVCKDGRDSNGLNKTIVVVCYKTPGNLNPPIYQGGTTCA--ATCSGCPSGISCTRS 202  
 QY 213 -NLG 215  
 Db 203 TGLC 206

## RESULT 9

T19852

hypoetical protein C39B9.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004



A:Molecule type: mRNA  
A:Residues: 1-243 <ME>  
A:Cross-references: UNIPROT:O8A205; UNIPARC:UPI00000E76C9; DBPJ\_AB009662; NID:G3374579;  
C:Comment: This protein functions as a cell adhesion protein for the association between  
C:GeneID: 81  
A:Map position: 17  
C:Superfamily: cysteine-rich secretory protein 1

Query Match	9.5%;	Score 119;	DB 2;	Length 243;
Best Local Similarity	21.6%;	Pred. No. 0.0044;		
Matches	42;	Conservative	29;	Mismatches 85;
			Indels	38;
			Gaps	8

```

Qy 41 HDVARNALATGGLINMTVNADAVILGPQNNMYKVMDCNLEBVAQQLAPC-----ND 94
Db 45 HNELRRQVS-----PPRSNLIKHEMNVOAANAQAQKANNCLIEHSSTE 87
Qy 95 PLPINTSLAONIARWLYFKDSEBEETVLQGVSWYMWYSASLGFMKGTLLDQFANQMAEPLAN 154
Db 88 DRKINIKKQGENL-YMSTDPTSMRTYLQ-SWYEEHNFVFGGALPNSAVGHYTG---- 139
Qy 155 IANTRNKRYGCAHKICPAQONNV-----SCYGSFKLAPNBYIQEGKACV-CDARPPS 208
Db 140 LVMYSSFRKVGCCVAYCENQDTLKYFYVCHYCEMGN-NVMKKSTPYHQITPCASCPCNNCDN 198
Qy 209 FCCDNLCDTRDAAS 222
Db 199 GLCTNSCDPEDLLS 212

```

### RESULT 13

hypothetical protein C39E.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: J119849  
R:Sim, M.  
Submitted to the EMBL Data Library, March 1996  
A:Reference number: Z19187  
A:Accession: J119849  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-210 <M11>  
A:Cross-references: UNIPROT:Q16540; UNIPARC:UPI0000082735; EMBL:Z70307; PDB:CAA4332.1  
A:Experimental source: clone C39E9  
C:Genetics:  
A:Gene: C39P.C39E.6  
A:Map position: 4  
A:Insertions: 58/5, 178/2  
A:Superfamily: yellowjacket venom allergen anti5en 5

Query Match	9.3%;	Score 117;	DB 2;	Length 210;
Best Local Similarity	22.4%;	Pred. No. 0.0056;		
Matches 57;	Conservative 32;	Mismatches 80;	Indels 86;	Gaps 13;

QY 1 MSAVVAVVALLALSYAEAGCCPNLSLQ--SDSARQIFLDPHNDRENTALGNLLIMWT 59  
Db 1 MRLTLLAV-----ACGVYTAQSGEGKOSTLNAHNDIRRIAGNTVAAKGR 48  
QY 60 NADAVILGPANMYKVDMDCNLEBVAQAQIAPCNDPLPIINTSLAQNIAFMLYFQDS--EEB 118  
Db 49 KESAT-----NMKKMKDSSLQ-----SAQNYANGCHMGHSTNDK 84  
QY 119 TVLQGVSMVYVSAALGFMKGTIKLDOF-----ANQMAELIAN----- 154  
Db 85 TIGENLITWESGDGF-----SDLDFGKIATVANDHEFEOGMMNSNKFSLALFPTGYAHA 139  
QY 155 --IANYNRKVGCAHKICPAQO-----NMVVS CVYSGPKLAPN---EVIWQEGKACVCD 203  
Db 140 TQIAMAFTGKIGCGVKNCGARDARRGGLFQVALYVQY---RVRNGFFPKINYNISG--ATCS 194  
QY 204 ARPDSPCCD---NLG 215  
Db 195 ACPAGTSCBQSTGLC 209

RESULT 14

hypothetical protein P49E11.5 - *Caenorhabditis elegans*  
 C|Species: *Caenorhabditis elegans*  
 C|Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C|Accession: T22437

A:Submitted to the EMBL Data Library, March 1996  
A:Reference number: Z19564  
A:Accession: T22437  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-212 <Wild>  
A:Cross-references: UNIPROT:Q20609; UNIPARC:UP1000008161B; EMBL:Z70308; PIDD:CAA54349.1  
A:Experimental source: clone P49B11  
C:Genetics:  
A:Gene: CBSP:P49B11.5  
A:Map position: 4  
A:Introns: 17/1; 146/3; 180/2  
C:Superfamily: yellowjacket venom allergen antigen 5

Query Match	9.3%	Score 116.5;	DB 2;	Length 212;
Best Local Similarity	24.3%;	Pred. No. 0.0063;		
Matches 55;	Conservative 32;	Mismatches 98;	Indels 41;	Gaps 13

QY 9 VLLALTSYAEAGCCCNLSOSDSARQIFLPHNDVRNIMLNGNLINTVNDADAVILGP 68  
 Db 8 ICLILTSF-----CETLCFSEBTGKNTILSRHNTLRSQIMLGYVAGNSNTRBS----- 55  
 QY 69 AONMYTVDWDCNLBEVAAQALPACNDPLPINTSLAQ-NIARMLYF-----KDSBEETVL 121  
 Db 56 ASNMKMLIDWTLLETTAQQDYSTGC---PTGHSASRANIGENMTWMTSPVYQTQDALLG 111  
 QY 122 QQVSTWTVS--ASLGFMTKTKLDQFANQMAEPLANIANIYRRKRYCGCAHKICPAQO--NM 176  
 Db 112 NRSANIMBEBEFPFGNGNLTLEBIFNSGIGHATQMAWATTNKGICGSIKSDSPSTGYOY 171  
 QY 177 VVSCTVYGPKLAPNEY---IMWEGKACY-CDAPRPSFCOM---LC 215  
 Db 172 VVVCILY-SP--AGNYIGMDIYKSGTGNC---PDGTNCBSSTGLC 211

## RESULT 15

hypothetical protein T05A10.5 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T24493

submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19898  
A:Accession: T24493  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-246 <Wtl>  
A:Cross-references: UNIPROT:P90958; UNIPARC:UP10000007AFB8; EMBL:Z68108; PIDD:CAA92136.1;  
A:Experimental source: Clone T05A10  
C:Genetics:  
A:Gene: CESP.T05A10..5  
A:Map position: X  
A:Initrns: 66/2, 142/3, 179/3, 208/2  
A:Superfamily: cysteine-rich secretory protein 1

Query Match	9.2%	Score 115.5;	DB 2;	Length 246;
Best Local Similarity	25.2%	Pred. No. 0.0091;		
Matches	58;	Conservative	27;	Mismatches 80;
				Indels 65;
				Gaps 14

Oy 21 FCCPNSLSGSDSARQIFLDFHNDVRRRIALGNGLIINTVNADAVILGP-AQNMRYKVDWC 79  
 | | | | | : | | | | | : | | | | |  
 Db 43 FCCDNL-VSDVTRNFTLBOGNFYRSRLAKG---FEW--NGETNTSQPKASQMIKMEYDC 96

```

Qy 80 NLEEVAAQIAPC-----NDPLP---INTSLAQNIARMLYKQD 114
Db 97 MLERFAQNMANNCFPAHSAHYERPNOGNTLYMSFSFNDPRLIHTA---VERK--WQE 150
Qy 115 SEB-----ETVLQGVSWTWVSASLGFMKGTLDQFANQMAEPLANIANYRNRKVCAHFI 169
Db 151 LEEFGTPIDNVLTPELW-----DLGKALIGHY-----TQMAWDRTRYRLGGIYN 194
Qy 170 CPAQNMVWSCVYGSFKLAENEVIWQBGKAC-VCDARPDFCCD---NLG 215
Db 195 CPKMSYVV--CHYGPAGNRKNKIYEIGDPCEVDDDCPIGTDCBKTSILC 242

```

Search completed: March 3, 2006, 19:10:13  
 Job time : 41 secs



## ORIGIN

RRNIALGNGLIINMTVNADAVILGPANNNYKVDMDNCNLEBVAOQIAPCNPLPINTSL  
AONTARMLYFDSEETVLQOVSWYWSASIGFMKGTLDGPANQMAPPLANIAYRN  
RKVGCAHRIKCPAQOONMNVSCVYGSPLKLPNEVIMQEKACVCDARPDFCCDNLCDDR  
DASVRHOCASP"

Query Match 100.0%; Score 763; DB 2; Length 763;  
Best Local Similarity 100.0%; Pred. No. 4.5e-226;  
Matches 763; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAGAACTGCTATGTCGGCGGCTGTGTGAGTTCCTGCTCTCGGCCCTGTTCTCTATGC 60
DB 1 GAGAACTGCTATGTCGGCGGCTGTGTGAGTTCCTGCTCTCGGCCCTGTTCTCTATGC 60
QY 61 CGAAGCAGGCTTTTGTGTCCGAAATAGTAAAGCCAAAGTGAACAGCGGAGGCAATTTT 120
DB 61 CGAAGCAGGCTTTTGTGTCCGAAATAGTAAAGCCAAAGTGAACAGCGGAGGCAATTTT 120
QY 121 CCTCGATTTTCACAATGATGTCGTGCAAAATATAGCACTTGGAAATGTTGATTAAC 180
DB 121 CCTCGATTTTCACAATGATGTCGTGCAAAATATAGCACTTGGAAATGTTGATTAAC 180
QY 181 GACAGTAAATGCAAGCGCGTCAATTTTGTGCTCAAGCTCAAGAAATGTAAGTGA 240
DB 181 GACAGTAAATGCAAGCGCGTCAATTTTGTGCTCAAGCTCAAGAAATGTAAGTGA 240
QY 241 GGAATGCACTTGGAAAGTAGACCAACAGATTTGGCCATGCAATGATCCCTAC 300
DB 241 GGAATGCACTTGGAAAGTAGACCAACAGATTTGGCCATGCAATGATCCCTAC 300
QY 301 GATTAATACCAAGCTGTGCTCAAAATATGCTAGATGCTGTACTTCAAGACAGTGA 360
DB 301 GATTAATACCAAGCTGTGCTCAAAATATGCTAGATGCTGTACTTCAAGACAGTGA 360
QY 361 AGAGACAGTTCTGCAACAAGTATCGTATGAGTGGTGAACGCGATCGCTGGATTATGA 420
DB 361 AGAGACAGTTCTGCAACAAGTATCGTATGAGTGGTGAACGCGATCGCTGGATTATGA 420
QY 421 AGGCAAGAACTTGAACCAATTTTCTTAACAGATGGGCTGAACCTTGAAGCAATTTG 480
DB 421 AGGCAAGAACTTGAACCAATTTTCTTAACAGATGGGCTGAACCTTGAAGCAATTTG 480
QY 481 CTATAGAAACCGAAAGGTTGATGTCATTAAGATCTGCCGCTCAGCAAAACATG 540
DB 481 CTATAGAAACCGAAAGGTTGATGTCATTAAGATCTGCCGCTCAGCAAAACATG 540
QY 541 AGTATCTGCTGTATGAAAGCCCAAACTTGAACCGAAAGATTATCTGCAAGAG 600
DB 541 AGTATCTGCTGTATGAAAGCCCAAACTTGAACCGAAAGATTATCTGCAAGAG 600
QY 601 AAAAGCTTGTGTGTGAGAGCTGTGTCATTAATCTCTGCGAACAACCTGTGAC 660
DB 601 AAAAGCTTGTGTGTGAGAGCTGTGTCATTAATCTCTGCGAACAACCTGTGAC 660
QY 661 GCGAGATGCTGCAAGTGTGCGCAACAGTGTGCGGCTGCGCATGAAGGAAAGAA 720
DB 661 GCGAGATGCTGCAAGTGTGCGCAACAGTGTGCGGCTGCGCATGAAGGAAAGAA 720
QY 721 GGTAGTCACTCCGCAATTAATATTCATGCAAAAAA 763
DB 721 GGTAGTCACTCCGCAATTAATATTCATGCAAAAAA 763

```

RESULT 2  
LOCUS CQ794479 763 bp DNA linear PAT 19-APR-2004  
DEFINITION Sequence 9 from Patent WO2004024769.  
ACCESSION CQ794479  
VERSION CQ794479.1 GI:46407112

KEYWORDS  
SOURCE Ostertagia ostertagi  
ORGANISM Ostertagia ostertagi

Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;

REFERENCE  
1 Trichostrongylidae; Haemonchidae; Ostertaginae; Ostertagia.  
Gelhof, P., Vercauteren, I., Claerebout, E., Vercruyse, J. and de  
Maere, V.  
TITLE Ostertagia vaccine  
JOURNAL Patent: WO 2004024769-A 9 25-MAR-2004;  
UNIVERSITEIT GENT (BE)  
FEATURES  
source Location/Qualifiers  
1..763  
/organism="Ostertagia ostertagi"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:6317"  
11..706  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CAG26297.1"  
/db\_xref="GI:46407113"

## CDS

```

QY 1 GAGAACTGCTATGTCGGCGGCTGTGTGAGTTCCTGCTCTCGGCCCTGTTCTCTATGC 60
DB 1 GAGAACTGCTATGTCGGCGGCTGTGTGAGTTCCTGCTCTCGGCCCTGTTCTCTATGC 60
QY 61 CGAAGCAGGCTTTTGTGTCCGAAATAGTAAAGCCAAAGTGAACAGCGGAGGCAATTTT 120
DB 61 CGAAGCAGGCTTTTGTGTCCGAAATAGTAAAGCCAAAGTGAACAGCGGAGGCAATTTT 120
QY 121 CCTCGATTTTCACAATGATGTCGTGCAAAATATAGCACTTGGAAATGTTGATTAAC 180
DB 121 CCTCGATTTTCACAATGATGTCGTGCAAAATATAGCACTTGGAAATGTTGATTAAC 180
QY 181 GACAGTAAATGCAAGCGCGTCAATTTTGTGCTCAAGCTCAAGAAATGTAAGTGA 240
DB 181 GACAGTAAATGCAAGCGCGTCAATTTTGTGCTCAAGCTCAAGAAATGTAAGTGA 240
QY 241 GGAATGCACTTGGAAAGTAGACCAACAGATTTGGCCATGCAATGATCCCTAC 300
DB 241 GGAATGCACTTGGAAAGTAGACCAACAGATTTGGCCATGCAATGATCCCTAC 300
QY 301 GATTAATACCAAGCTGTGCTCAAAATATGCTAGATGCTGTACTTCAAGACAGTGA 360
DB 301 GATTAATACCAAGCTGTGCTCAAAATATGCTAGATGCTGTACTTCAAGACAGTGA 360
QY 361 AGAGACAGTTCTGCAACAAGTATCGTATGAGTGGTGAACGCGATCGCTGGATTATGA 420
DB 361 AGAGACAGTTCTGCAACAAGTATCGTATGAGTGGTGAACGCGATCGCTGGATTATGA 420
QY 421 AGGCAAGAACTTGAACCAATTTTCTTAACAGATGGGCTGAACCTTGAAGCAATTTG 480
DB 421 AGGCAAGAACTTGAACCAATTTTCTTAACAGATGGGCTGAACCTTGAAGCAATTTG 480
QY 481 CTATAGAAACCGAAAGGTTGATGTCATTAAGATCTGCCGCTCAGCAAAACATG 540
DB 481 CTATAGAAACCGAAAGGTTGATGTCATTAAGATCTGCCGCTCAGCAAAACATG 540
QY 541 AGTATCTGCTGTATGAAAGCCCAAACTTGAACCGAAAGATTATCTGCAAGAG 600
DB 541 AGTATCTGCTGTATGAAAGCCCAAACTTGAACCGAAAGATTATCTGCAAGAG 600
QY 601 AAAAGCTTGTGTGTGAGAGCTGTGTCATTAATCTCTGCGAACAACCTGTGAC 660
DB 601 AAAAGCTTGTGTGTGAGAGCTGTGTCATTAATCTCTGCGAACAACCTGTGAC 660
QY 661 GCGAGATGCTGCAAGTGTGCGCAACAGTGTGCGGCTGCGCATGAAGGAAAGAA 720

```



```

Db      661 GCGAGATGCTGCGACGATGTCGCCACGATGCGGTCGCCACGATGAAGCGAAAGAAAT 720
QY      721 GGTAGTCAACCCCGAATTAATATTCATGCAAAAAA 763
Db      721 GGTAGTCAACCCCGAATTAATATTCATGCAAAAAA 763

RESULT 3
LOCUS   CO794471 828 bp DNA linear PAT 19-APR-2004
DEFINITION Sequence 1 from Patent WO2004024769.
ACCESSION CO794471
VERSION   CO794471.1 GI:46407104
KEYWORDS
SOURCE    Osterlagia osterlagi
ORGANISM  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
           Trichostrongyloidea; Haemonchidae; Osterlagiinae; Osterlagia.
REFERENCE 1 Geldhof, P., Vercauteren, I., Claerebout, E., Vercruyse, J. and de
AUTHORS   Maere, V.
TITLE      Osterlagia vaccine
JOURNAL    Patent: WO 2004024769-A 1 25-MAR-2004;
           UNIVERSITEIT GENT (BE)
FEATURES   Location/Qualifiers
            1..828
             /organism="Osterlagia osterlagi"
             /mol_type="unassigned DNA"
             /db_xref="taxon:6317"
             11..721
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="CAG26291.1"
              /db_xref="GI:46407105"
              /translation="MQALIGIALYLVLTSTNTEAGPCCPADLNTDARXIFLDFHN
              OVRDLAGASPLNLTGAVQMRVLTGAPARKRMMDCLAKAKMIPCTPLPID
              TSIPLQALQMLFQNSQKXEVLTQTPMSVTLTLRLQDTEANTYNNQIRPLSNAN
              WQNIKVGCAHKYCKKPTKPTGNMVSCHVGEVLDQNEVMDKPTCMANAYRSPFCNN
              LCDTIAATLRQPKCKST"

ORIGIN
Query Match 35.8%; Score 273.4; DB 6; Length 828;
Best Local Similarity 63.6%; Pred. No. 1.9e-73;
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2;

QY      38 CTCCTGGCCCTGTTCTCTATGCGGAGAGAGGCTTTGTGTCGAATAGTCAACCA 97
Db      44 CTGGTCTGCTGATCAATACCGAAGAGGTTTGTGCGCCACAGATCTAACCAG 103
QY      98 AGTGACAGGCGGAGAGATTTTCTCGATTTTCAAGATGTTGCGAATATAGCA 157
Db      104 ACTGATGAGGCAAGAAATCTTCTCGATTTTCAAGATGTTGCGCGTGAATAGCA 163
QY      158 CTGGAAATGTTGATTAAGTGAAGAGAGATTAATGACAGCGGCTCA---TTCCTGGTCA 214
Db      164 GGTGACAGCCGCTGCTCAACTCAGCGAGGCTTTCAATGAGGAATGTTCTCGTCA 223
QY      215 GCTCAAGACATGTAAGATGAGCTGCGAATGTCGAAGAGTACAGACAGCAACAG 274
Db      224 GCTAAGACATGTAAGATGAGCTGCGAATGTCGAAGAGTACAGACAGCAACAG 283
QY      275 ATTGGCCATGCAATGATCCCTTACCGATTAATCAAGCTGCTCAAAATATCGCTAGA 334
Db      284 ATTGGCCATGCAATGATCCCTTACCGATTAATCAAGCTGCTCAAAATATCGCTCAR 343
QY      335 TGGCTGTAATCTTAAGAGAGTGAAGAGAGAGAGAGAGTCTGCAACAGTATCGGTATG 394
Db      344 TGGCTGTAATCTTAAGAGAGTGAAGAGAGAGAGAGAGTCTGCAACAGTATCGGTATG 403
QY      395 GTAGAGCATGCTGCGATTTATGAAGAGCAGAACTTGAACCAATTTGCTAACGATG 454
Db      404 GTAGAGCATGCTGCGATTTATGAAGAGCAGAACTTGAACCAATTTATTAATCGGCA 463

```

```

QY      455 GCTGAACCTCTAGCAAACTTGCAACTATAGAAACCGAAAGTTGATGTCCTAAG 514
Db      464 ATTAGACCACTATTCACACTTGGAAGTGCAGAAACCTTAAGTTGATGTCACAA 523
QY      515 ATCTG-----CCCGCTCAGCAAAACATGTAATCTCGGTATGGAAGCCCA 568
Db      524 GTGTGCAAAATCCCGACCGGACAAATATGTTGTCTTGCGCTATGCGCGCAAGTA 583
QY      569 CTTCACCGAAGCAATTTATCTGCGAGAGAGAAAGCTTTGTGTGCGACGTCCTCA 628
Db      584 CTTCAGATAGCAAGATTTATGAGCAAGAGCAACCTTGCAATGCAATGCTTATCC 643
QY      629 GATTCAATCTGCTGCGACAACTGTTGACACCGAGATGTCGAGTGTGCGCACAG 688
Db      644 AACTGTTCTGCTGCAACATCTGTGACACAAATGCTGTGCGACACTTGCGACAG 703
QY      689 TGTTCGCGCTGCGCATGAAGCGAAAGAAATTTGATGCAACC 731
Db      704 CCTTGAAATCGACTTGAAGCGAAAGGCGTTGTGATGTCCTCC 746

RESULT 4
LOCUS   OOS310812 712 bp mRNA linear INV 12-MAR-2003
DEFINITION Osterlagia osterlagi mRNA for ancylostoma-secreted protein-1like
           protein (OSP-2 gene), clone Fl75.
ACCESSION AJ310812
VERSION   AJ310812.2 GI:18860823
KEYWORDS  ancylostoma-secreted protein-1like protein; OSP-2 gene.
SOURCE    Osterlagia osterlagi
ORGANISM  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
           Trichostrongyloidea; Haemonchidae; Osterlagiinae; Osterlagia.
REFERENCE 1 Geldhof, P., Peelaers, I., Claerebout, E., Berx, G. and
AUTHORS   Vercruyse, J.
TITLE      Identification of excretory-secretory products of larval and adult
           Osterlagia osterlagi by immunoscreening of cDNA libraries
           Mol. Biochem. Parasitol. 126 (2), 201-208 (2003)
           12615319
JOURNAL   Parasitol. 126 (2), 201-208 (2003)
PUBMED    12615319
REFERENCE 2 Vercauteren, I.J.R.
AUTHORS   Vercauteren, I.J.R.
TITLE      Direct Submission
JOURNAL   Submitted (29-MAR-2001) Vercauteren I.J.R., Department of
           Parasitology, Ghent University, Faculty of Veterinary Medicine,
           Salisburylaan, 133, B-9820 MERELBEKE, BELGIUM
           Revised by [3]
REFERENCE 3 Vercauteren, I.J.R.
AUTHORS   Vercauteren, I.J.R.
TITLE      Direct Submission
JOURNAL   Submitted (13-AUG-2001) Vercauteren I.J.R., Department of
           Parasitology, Ghent University, Faculty of Veterinary Medicine,
           Salisburylaan, 133, B-9820 MERELBEKE, BELGIUM
REMARK    4 (bases 1 to 712)
REFERENCE 4 Vercauteren, I.J.R.
AUTHORS   Vercauteren, I.J.R.
TITLE      Direct Submission
JOURNAL   Submitted (07-FEB-2002) Vercauteren I.J.R., Department of
           Parasitology, Ghent University, Faculty of Veterinary Medicine,
           Salisburylaan, 133, B-9820 MERELBEKE, BELGIUM
           On Feb 22, 2002 this sequence version replaced gi:13536973.
COMMENT   Location/Qualifiers
           1..712
            /organism="Osterlagia osterlagi"
            /mol_type="mRNA"
            /db_xref="taxon:6317"
            /clone="Fl75"
            /dev_stage="adult"
            1..712
             /gene="OSP-2"
             <1..711
              /gene="OSP-2"

```

/codon\_start=1  
/product="ancylostoma-secreted protein-like protein"  
/protein\_id="CAD2183.1"  
/db\_xref="GI:18860824"  
/db\_xref="GOA:O8T8P4"  
/db\_xref="UniProt/TREMBL:O8T8P4"  
/translation="MQLIGTALYIVTSTNAGPCCPADLNQDEARXIPLDPHN  
QVRDITAGASPLNLTGAIVQKENVIGPKAKMYRDNDCNLEAKKAMTPTPLPID  
TSIPNLQMLFQNSQKXEVLTPTPMSVSLANLPDTEATYIMQIPLSLIAN  
MQLKVGCAHYCKEPTGTNNVSCAYGSEVLQDNEVYMDKPTCMCAVNSFCNN  
LQDITAAITLRNPKST"

## ORIGIN

Query Match 34.6%; Score 263.8; DB 2; Length 712;  
Best Local Similarity 63.6%; Pred. No. 1.8e-70;  
Matches 432; Conservative 3; Mismatches 235; Indels 9; Gaps 2;

QY 38 CTCTGCGCCCTGTTCTCCATGCGGAGCGCTTTGTGTCCGATAGTCTAAGCAA 97  
DB 34 CTGCTGCTGTGACATCAATACCAAGAGGCTTTTGTGCTGCCGACGATCTTAACCA 93  
QY 98 AGTGCAGGCGAGGACAGATTTTCTCGATTTTCACAATGATGTGTGCAAAATATAGCA 157  
DB 94 ACTGATGAGGCAAGAAAATCTTCTCGATTTTCACAATGATGTGTGCAAAATATAGCA 153  
QY 158 CTGGAATATGTTTATTAATGAGACATGAAATGACAGCGGCTCA---TTCTGTGCTCA 214  
DB 154 GGTGCAAGCCGCTGCTCAACCTCACCGAGCTGTTCATATCGAATGTTCTGATCA 213  
QY 215 GCTCAAGATATTAAGATGAGCTGGATGCACTTGGAGAAGATAGCAGACAGCAAG 274  
DB 214 GCTAAGACATATGAGAAATGAGCTGGATGCACTTGGAGAAGATAGCAGACAGCAAG 273  
QY 275 ATTGCGCCATGAGATGATCCCTACCGATTAATACCAAGCTGCTCAAAATATGCTAGA 334  
DB 274 ATTGCGCCATGAGATGATCCCTACCGATTAATACCAAGCTGCTCAAAATATGCTAGA 333  
QY 335 TGGCTGTACTTCAAGACATGAGAGAGACAGATTTCTGCAACAGTATGCTGATTTGG 394  
DB 334 TGGCTGTACTTCAAGACATGAGAGAGAGATTTCTGCAACAGTATGCTGATTTGG 393  
QY 395 GTGAGGCGATGCTGGATTTATGAAAGGACAGAACTTGACCAATTTGCTAACAGTGG 454  
DB 394 GTAAACGCGATCACTAGAAATCTTCACTGATACAGAACTTAATTAATGCTGCAAA 453  
QY 455 GCTGAACCTCTAGCAAACTTGCAAACTATAGAAACCGAAAGTTGATGATGCTCCATAAG 514  
DB 454 ATTGACCACTATCAACATTTGGAATCTGGCAAACTTAATTAATGATGCTGCAAA 513  
QY 515 ATCTG-----CCCGCTCAGCAAAACATGATGATCTGCTGATGATGAGAGCCCAAA 568  
DB 514 GTGTGCAAAATCCCAACCGGGAACAAATATGATGATGCTGCTGATGATGAGAGAG 573  
QY 569 CTTCACACCAAGCAAGTATCTGGCAGAGAAAGGCTGTGTGTCGACGCTGCTCA 628  
DB 574 CTCAAGATACCAAGTGTATGAGCAAGAGGACCACTTGCAATGTCATGCTTAATCC 633  
QY 629 GATTCAATCTGCTGCGCAACTGTGTGACACCGGAGATGCTGCAATGCTGCTGCAACAG 688  
DB 634 AATCTGCTGCTGCGCAACTGTGTGACACCAATGCTGCTGCAACTGCTGCAACAG 693  
QY 689 TGTGCGCGTGCCTATGAA 707  
DB 694 CTTGCAATCGACTTGA 712

RESULT 5  
AX655393 2000 bp DNA linear PAT 22-MAR-2003  
LOCUS AX655393  
DEFINITION Sequence 5263 from Patent WO03000898.  
ACCESSION AX655393  
VERSION AX655393.1 GI:29158207  
KEYWORDS

## SOURCE

Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Echinozoidae; Oryzaceae; Oryza.

## REFERENCE

AUTHORS Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
Katafiri,F., Quan,S., Tao,Y., Whitham,S., Xie,Z., Zhu,T. and Zou,G.  
TITLE Plant genes involved in defense against pathogens  
JOURNAL Patent: WO 03000898-A 5263 03-JAN-2003;  
SynGene PatCipations AG (CH)

## FEATURES

source

1..2000  
Location/Qualifiers  
/organism="Oryza sativa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:4530"

## ORIGIN

Query Match 5.4%; Score 41.4; DB 6; Length 2000;  
Best Local Similarity 9.3%; Pred. No. 0.4;  
Matches 54; Conservative 263; Mismatches 264; Indels 2; Gaps 1;

QY 167 GCTTGAATACTGACAGTAAATGACAGCGGCTCATTTCTGTCACGCTCAAGCAATG 226  
DB 219 RKGTYKMTTYSASRCWRAYMTTYSWACSSYTWCRKSMWKKRKRKRSYGYW 278  
QY 227 TACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 286  
DB 279 SMSYMMWCTAYKYSYRWCTYRGSGWRATYWGSGWRSBAMWTKMYWRYGYKW 338  
QY 287 AATGATCCCTACCATTAATATACAGCCTGCTCAAAATATGCTAGATGCTGACTTC 346  
DB 339 KRGWAGGMMWRSMWRMSKACYWRWRWMTTRRRWAKKSRTSRKRKRKWKRYK 398  
QY 347 AAAGCACTGAAGAGAGACAGTCTGCAACAAGTATGCTGATTTGGTGAAGCCATCG 406  
DB 399 RMRGYSRMRSCRRARWMRCSRGRAMWRCRCRCMTCRKYSYGMWKKWKRRASTYKMS 458  
QY 407 CTGGATTTATGAAGGACCAACTTGACCAATTTGCTAACAGTGGGCTGAACCTCTA 466  
DB 459 RMYRKRKKCSRTTWGTRGMMGTWRCRYKRSKMKRCRRRMRMRMYRMYRMYR 518  
QY 467 GCAACATTTGCAAACTATAGAAACCAAGGTTGG--ATGCGCCATAGATCTGCCCG 524  
DB 519 SARMYRRCARKKYSYSAARARACMYRGGYTMAAGMMKRYRMYRMYRMYRMYR 578  
QY 525 CTCAGCAAAACATGCTAGTATCTCGCTGATGAGAGCCCAACTTGACCGAAGCAAG 584  
DB 579 SWYCRMSYASCMKARARAGAMCRSRKMSKMSKRCRCASKRSSAKRYAMWG 638  
QY 585 TTATCTGCGAGAGAAAGGAGCTGTGTGTCGACGCTGTCAGATTCATTTCTGCTGCG 644  
DB 639 MTSGSRMSRMSYTYCYWRKMSKSTCTMYTMSKTYTAKYSYRMYRMYRMYRMYR 698  
QY 645 ACAACCTGTGACAGCGAGATGCTGCGAGTGTGCGACCAAGTGTGCGGCTGCCAT 704  
DB 699 YRYSYMYTMYAMYSSTMTAMTMKYSGRYMTYSYKCYCKGSKYRSMYYSWMAKTM 758  
QY 705 GAAGCGAAGAAATTTGATGACCCCGCAATAATATTCAT 747  
DB 759 KMRRYATRMMWMYRYSKMYTCTMGMGYWYMYRMYRMYR 801

RESULT 6  
AP008208 049/c  
WPCOMMENT  
Sequence split into 360 fragments LOCUS AP008208 Accession AP008208  
Fragment Name Begin End  
AP008208\_000 1 110000  
AP008208\_001 10001 210000  
AP008208\_002 20001 310000  
AP008208\_003 30001 410000  
AP008208\_004 40001 510000

AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000
AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000
AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000

AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	10110000
AP008208_101	10100001	10210000
AP008208_102	10200001	10310000
AP008208_103	10300001	10410000
AP008208_104	10400001	10510000
AP008208_105	10500001	10610000
AP008208_106	10600001	10710000
AP008208_107	10700001	10810000
AP008208_108	10800001	10910000
AP008208_109	10900001	11010000
AP008208_110	11000001	11110000
AP008208_111	11100001	11210000
AP008208_112	11200001	11310000
AP008208_113	11300001	11410000
AP008208_114	11400001	11510000
AP008208_115	11500001	11610000
AP008208_116	11600001	11710000
AP008208_117	11700001	11810000
AP008208_118	11800001	11910000
AP008208_119	11900001	12010000
AP008208_120	12000001	12110000
AP008208_121	12100001	12210000
AP008208_122	12200001	12310000
AP008208_123	12300001	12410000
AP008208_124	12400001	12510000
AP008208_125	12500001	12610000
AP008208_126	12600001	12710000
AP008208_127	12700001	12810000
AP008208_128	12800001	12910000
AP008208_129	12900001	13010000
AP008208_130	13000001	13110000
AP008208_131	13100001	13210000
AP008208_132	13200001	13310000
AP008208_133	13300001	13410000
AP008208_134	13400001	13510000
AP008208_135	13500001	13610000
AP008208_136	13600001	13710000
AP008208_137	13700001	13810000
AP008208_138	13800001	13910000
AP008208_139	13900001	14010000
AP008208_140	14000001	14110000
AP008208_141	14100001	14210000
AP008208_142	14200001	14310000
AP008208_143	14300001	14410000
AP008208_144	14400001	14510000
AP008208_145	14500001	14610000
AP008208_146	14600001	14710000
AP008208_147	14700001	14810000
AP008208_148	14800001	14910000
AP008208_149	14900001	15010000
AP008208_150	15000001	15110000

```

AP008208_151 15100001 15210000
AP008208_152 15200001 15310000
AP008208_153 15300001 15410000
AP008208_154 15400001 15510000
AP008208_155 15500001 15610000
AP008208_156 15600001 15710000
AP008208_157 15700001 15810000
AP008208_158 15800001 15910000
AP008208_159 15900001 16010000
AP008208_160 16000001 16110000
AP008208_161 16100001 16210000
AP008208_162 16200001 16310000
AP008208_163 16300001 16410000
AP008208_164 16400001 16510000
AP008208_165 16500001 16610000
AP008208_166 16600001 16710000
AP008208_167 16700001 16810000
AP008208_168 16800001 16910000
AP008208_169 16900001 17010000
AP008208_170 17000001 17110000
AP008208_171 17100001 17210000
AP008208_172 17200001 17310000
AP008208_173 17300001 17410000
AP008208_174 17400001 17510000
AP008208_175 17500001 17610000
AP008208_176 17600001 17710000
AP008208_177 17700001 17810000
AP008208_178 17800001 17910000
AP008208_179 17900001 18010000
AP008208_180 18000001 18110000
AP008208_181 18100001 18210000
AP008208_182 18200001 18310000
AP008208_183 18300001 18410000
AP008208_184 18400001 18510000
AP008208_185 18500001 18610000
AP008208_186 18600001 18710000
AP008208_187 18700001 18810000
AP008208_188 18800001 18910000
AP008208_189 18900001 19010000
AP008208_190 19000001 19110000
AP008208_191 19100001 19210000
AP008208_192 19200001 19310000
AP008208_193 19300001 19410000
AP008208_194 19400001 19510000
AP008208_195 19500001 19610000
AP008208_196 19600001 19710000

Query Match 5.4%; Score 41.4; DB 15; Length 110000;
Best Local Similarity 54.2%; Pred. No. 0.47; 71; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 360 AAGAGACAGTTCGCAACAAGTATCGTGATTTGGTGAGCGCATCGCTGGATTATTGA 419
DB 88566 AAGTGAAGTTCGATGATGATGCTCCCTTACATGAGTGAATAGATCGCAAGCAAGA 88507
QY 420 AAGGACGAAACTTGAACCAATTTGCTAACCAAGTGGGCTGAACCTCTAGCAAACTTGCA 479
DB 88506 TTGGAGACAGAACTAATGATTTTACTACCAATTTGGCTTGAATTTTCACAAGATTGCA 88447
QY 480 ACTATAGAAACGGAAGGTTGATGTGCCCATAG 514
DB 88446 AATGTAAATTCAACATTTCTGGCTTCAAAATATG 88412

RESULT 7
LOCUS AP005849 132729 bp DNA linear PLN 16-NOV-2004
DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
ACCESSION AP005849
VERSION AP005849.3 GI:46390951
KEYWORDS
SOURCE Oryza sativa (japonica cultivar-group)

```

```

ORGANISM
Oryza sativa (japonica cultivar-group)
Bakeriotea; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriothalideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT
Oryza sativa (japonica cultivar-group)
Submitted (24-OCT-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki, 305-8602, Japan
(E-mail:tsasaki@nias.affrc.go.jp, URL:http://xgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Apr 14, 2004 this sequence version replaced gi:3450234.
Genes were predicted from the integrated results of the following:
GENSCAN (http://CCR-081.mt.edu/GENSCAN.html), GENESH
(http://www.softberry.com/), GeneMark.hmm
(http://opal.biology.gatech.edu/GeneMark/), Glimmer
(http://www.cigr.org/cdb/glimmer/glmr.form.html), RiceHMM
(http://xgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
(http://bioinformatics.laastate.edu/cgi-bin/sp.cgi), sim4
(http://globin.cse.psu.edu/html/docs/sim4.html), gap2
(http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
genomic sequence was searched against NCBI Nonredundant Protein
database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
sequence database at RGP or DBU. Protein homologues of the coding
regions were searched against NCBI Nonredundant Protein database
with BLASTP. ESTs represent the identified cDNA sequences using
BLASTN with the corresponding DBU accession no. and RGP clone ID.
Full-length cDNAs represent the identified cDNA sequences using
BLASTN with the corresponding DBU accession no.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with full-length cDNA or
EST homology (covering almost the entire length of partial
sequence) is classified as an 'unknown' protein. A gene predicted
by two or more gene prediction programs is classified as a
'hypothetical' protein according to IRGSP standard. A gene
predicted by a single gene prediction program is also classified as
a probable 'hypothetical' protein and is included as a
miscellaneous feature of the sequence.
The orientation of the sequence is from -21M3 to M13rev of the BAC
clone. This sequence of OSJNB0031B09 clone has an overlap with
OJ1145.F01 (DBU: AP004046) clone at 5' end and with P0453H10
(DBU: AP005785) clone at 3' end. Detailed information on overlap
and assembly quality together with annotation of this entry is
available at http://xgp.dna.affrc.go.jp/Genomeseq.html.

FEATURES
source
1..132729
location/Qualifiers
1..132729
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSJNB0031B09"
507..1259
/gene="OSJNB0031B09.1"
join(507..682,797..>1259)
/gene="OSJNB0031B09.1"
note="start and end point are not identified"
join(507..682,797..1259)
/gene="OSJNB0031B09.1"
note="predicted by GeneMark.hmm etc."
/codon_start=1
/product="hypothetical protein"
/protein_id="BAD16465.1"
/db_xref="GI:46390952"
/translation="MDARGARLTGDDGDDGDDVGARCGARRLRGGGAALCDGAAMW"

```

	misc_feature
gene	RARRRGVLAAMWGSGLVARGRADDDVTTRRGGAGGDAACGAAGRADGGTTGVR RRRGARGANRGGAAATATADGGCGSARBRASSTAAAPRPBGDSMAAABRRRN ROMAHGARAARASRSWAATAGSGCTGCAGYSARKMGVRLGNLVEHLADE" join(1242). .248,3888. .4800,6598. .6753,6862. .7095) /gene="OSJNB0031B09.2" join(1242). .248,3888. .4800,6598. .6753,6862. .7095) /gene="OSJNB0031B09.2" /note="copia-type polyprotein-like"
gene	/complement(7653. .8923) /gene="OSJNB0031B09.3" complement(join(7653). .7929,8479. .8923)) /gene="OSJNB0031B09.3" /note="supported by full-length cDNA(e) : AK120425"
CDS	complement(join(7821. .7929,8479. .8810)) /gene="OSJNB0031B09.3" /note="contains EST(e) : AUI61887(S6190),D39569(S1037) contains full-length cDNA(e) : AK120425"
gene	/codon start=1 /product="ribosomal protein S21-like protein"
mRNA	/protein_id="BAD16466.1" /db_xref="GI:46390953"
gene	/translation="MAAAPATTSLTALLQLPLAFPSGRSPPPSVVNAPPSPTA AVAAATGVYIVDNEGEDESIFPRFRFWRVGVLQEIQRKRRYESKKOERKRKAR BAARRNRREDESEPRPEPDDESAARARDDEKDMMIEIDGIL" complement(9908. .13128) /gene="OSJNB0031B09.4" complement(join(9908. .10162,11120. .11325,12777. .12893, 13018. .13128)) /gene="OSJNB0031B09.4" /note="supported by full-length cDNA(e) : AK111066"
CDS	complement(join(11149. .11325,12777. .12893,13018. .13071)) /gene="OSJNB0031B09.4" /note="contains EST(e) : AU092625(C52617),AU063035(C52617) contains full-length cDNA(e) : AK111066 similar to Arabidopsis thaliana chromosome 3, At3g01435"
mRNA	/codon start=1 /product="unknown protein"
gene	/protein_id="BAD16467.1" /db_xref="GI:46390954"
mRNA	/translation="MLPQSOSLSQRLNHVEOMTVRAVNLGTMBEIKGNATPRTEGA VAGHCRREFLMAKKEIQTLTBREIKSGCYRPFPEKDYARIANEICKKRYEVLEKDA AMOKNIEBSTGS"
gene	13885. .15378 /gene="OSJNB0031B09.5" join(13885. .14326,14395. .15378) /gene="OSJNB0031B09.5" /note="supported by full-length cDNA(e) : AK108567"
CDS	join(13898. .14326,14395. .15099) /gene="OSJNB0031B09.5" /note="contains full-length cDNA(e) : AK108567"
gene	/codon start=1 /product="putative Anter-specific proline-rich protein APG precursor"
mRNA	/protein_id="BAD16468.1" /db_xref="GI:46390955"
gene	/translation="MRSPRDLVAVALQLSLVLLLSGEPAAMAAGSKVAALIVFGD ETVDGNNNYISTLVSRDPFAPGRDLQLAGVGSGGGGRGTFGRFSNGRLAVPI SEAA FLPLPVLYADLVAVNMSSSLGACGFAAGGYGNATSDLPSTVPLMKEDVFEVYEA RLSRSGDDAAAATAATTSBALYIVTMGMNDLTENVYARBAHAAYSTAAAYGDY LLGVAARPYREIHALGAKRINDLNGIPPGICPIELRBATGACTEBRTNAVGRHAGLD MTARLNGELSGGARIVGDYGVGAVALVADLPAAIGVENVKAGCCGVGVGEMWGCA GARSPLCTGDSKAFMDAIHFHTERTLRADIAADAKNTTLHVFL"
gene	15811. .19509 /gene="OSJNB0031B09.6" join(15811. .16277,17884. .18365,19169. .19509) /gene="OSJNB0031B09.6" /note="supported by full-length cDNA(e) : AK060691"
mRNA	join(15882. .16277,17884. .18365,19169. .19365) /gene="OSJNB0031B09.6" /note="contains full-length cDNA(e) : AK060691"
CDS	/codon start=1 /product="putative gene-specific proline-rich protein APG precursor"

Query Match	Best Local Similarity	Match	84, Conservative	0; Mismatches	71; Indels	0; Gaps	0;
QY	360	AAGAGACAGTTTCGACCAAGTATCGGATGATGGGAGACGGATCGCTGGAGTTTATGCA	419				
DB	59692	AAGGTAAAGTTCAAGATGATGCTCTTCACATGAGTATGATGCGCAAGCAAAAGCA	59633				
QY	420	AAGCAGCAAACTTGACCAATTTGCTAACCAAGTGGGCTGAACCTTAGCAAACTTGC	479				
DB	59692	TTTGAGAGCAAGACTAATGATTTTACACATGATGGCTGTGAATTTTACAGAGATTGGCA	59577				
QY	480	ACTATAGAAACCGAAGGTTGATGGATGCGCCATAG	514				
DB	59572	AATGTGAATTCACATCTTCGGCTTCAATATG	59538				
RESULT 8							
LOCUS	AC131303	187110 bp	DNA	linear	ROD 02-SBF-2004		
DEFINITION	Mus musculus chromosome 1, clone RP24-119019, complete sequence.						
ACCESSION	AC131303						
VERSION	AC131303.8	GI:51854578					
KEYWORDS	HTG.						
SOURCE	Mus musculus (house mouse)						



```

repeat_region      /rpt_family="MTD"
complement(11921..12063)
/rpt_family="B1_MM"
repeat_region      complement(15576..16084)
/rpt_family="RLTR11A"
repeat_region      16235..16403
/rpt_family="RSINE1"
repeat_region      17073..17094
/rpt_family="AT_r1ch"
repeat_region      17106..17249
/rpt_family="B1_MM"
repeat_region      17504..17612
/rpt_family="(TTTC)n"
repeat_region      17637..17657
/rpt_family="AT_r1ch"
repeat_region      17692..18265
/rpt_family="L1_MM"
repeat_region      18337..18363
/rpt_family="AT_r1ch"
repeat_region      complement(18812..21567)
/rpt_family="L1VL2"
repeat_region      21576..21710
/rpt_family="(TTTC)n"
repeat_region      complement(21745..21880)
/rpt_family="B1_MM"
repeat_region      21886..22170
/rpt_family="MTD"
repeat_region      complement(22689..22899)
/rpt_family="B3"
repeat_region      22955..23104
/rpt_family="B1_MM"
repeat_region      23105..23124
/rpt_family="(A)n"
repeat_region      23523..23607
/rpt_family="MLT2B3"
repeat_region      23692..24043
/rpt_family="MLT2B3"
repeat_region      24990..25418
/rpt_family="L1_MM"
repeat_region      25443..25704
/rpt_family="L1_MM"
repeat_region      26072..26146
/rpt_family="(TCTCTG)n"
repeat_region      complement(26204..26233)
/rpt_family="B3A"
repeat_region      26234..26267
/rpt_family="(TA)n"
repeat_region      complement(26268..26415)

Query Match      5.4%; Score 41.2; DB 9; Length 187110;
Best Local Similarity 49.5%; Pzed. No. 0.56;
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

```

48 TGTTCCTATGCGGAGGAGGCTTTGTTGTCGATAGCTTAAGCCAAAGTGAAGCG 107  
 73210 TGAATTCACAGTGTGTAAGGCTATATCTGTGAAGTGTGACCTCAAGTAAAGAA 73269  
 108 CGAGGCAATTTTCTCGATTTTCAAAATGATTTGCGAAATATATAGCACTTGAAGT 167  
 73270 TTACAGAAATGTCATGCAATTTTATTTAAATAAAATCAAAAGACATTAAGCTTCAAGCAGTC 73329

168 GTTGTATAATCGAAGTAATAGCAAGCGGCTATTTGTCGAGCTCAGACGAAATAGT 227  
 73330 CTAAAGAAATCTGCAATAGTCAGAGGCGAGGGAAGCATTTTAAACTTAAACAAAATATA 73389  
 228 ACAAGTGAAGTGAAGTCAACTTGAAGAAAGT 261  
 73390 ACTCTGTAAGGCTGTGATAGTCTGACATTAAT 73423

Db 73390 ACTCTGTAAGGCTGTGATAGTCTGACATTAAT 73423

RESULT 9  
 AC123556/c  
 LOCUS AC123556 198900 bp DNA linear ROD 08-NOV-2003

```

DEFINITION      Mus musculus BAC clone RP23-176M9 from 1, complete sequence.
ACCESSION      AC123556
VERSION      AC123556.4
KEYWORDS      GI:29244829
SOURCE      HTG
ORGANISM      Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 198900)
Dauphin, S., Meyer, R., Kohlberg, S. and Haekenson, W.
The sequence of Mus musculus BAC clone RP23-176M9
Unpublished (2001)
2 (bases 1 to 198900)
Wilson, R.
Sequencing of Mus musculus
Unpublished (2001)
3 (bases 1 to 198900)
Mcpherson, J.D. and Waterston, R.H.
Direct Submision
Submitted (30-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 198900)
Mcpherson, J.D. and Waterston, R.H.
Direct Submision
Submitted (11-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 198900)
Mcpherson, J.D. and Waterston, R.H.
Direct Submision
Submitted (26-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 198900)
Wilson, R.
Direct Submision
Submitted (08-NOV-2003) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Mar 26, 2003 this sequence version replaced gi:28302047.
-----
Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@watson.wustl.edu
-----
Summary Statistics
Center project name: M_BA0176M09
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RP23-176 BAC library has been constructed by Kazuhiro Oseigawa and Minako Tateo in the laboratory of Pieter de Jong (<http://www.chori.org>) from female C57BL/6J mouse kidney and/or brain genomic DNA. The clone and detailed information can be obtained from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:  
This sequence is the entire insert of the clone. This clone is  
overlapped by AC117187.

FEATURES  
Source  
1..198900  
Location/Qualifiers

/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="1"  
/map="1"  
/clone="RP23-176M9"  
/clone\_1b="RPC1-23"  
2138..3200  
/rpt\_family="ERV1"  
repeat\_region  
3436..3494  
/rpt\_family="ERV1"  
repeat\_region  
4075..4320  
/rpt\_family="L1"  
repeat\_region  
4478..4549  
/rpt\_family="ERV1"  
repeat\_region  
4555..4803  
/rpt\_family="L1"  
repeat\_region  
4883..4937  
/rpt\_family="ERV1"  
repeat\_region  
4950..5305  
/rpt\_family="L1"  
repeat\_region  
5315..5367  
/rpt\_family="ERV1"  
repeat\_region  
5620..5822  
/rpt\_family="L1"  
repeat\_region  
6086..6363  
/rpt\_family="ERVK"  
repeat\_region  
6196..6437  
/rpt\_family="ERVK"  
repeat\_region  
6336..6616  
/rpt\_family="ERVK"  
repeat\_region  
6641..6691  
/rpt\_family="L1"  
repeat\_region  
7065..7180  
/rpt\_family="B4"  
repeat\_region  
7816..7956  
/rpt\_family="B4"  
repeat\_region  
12405..12486  
/rpt\_family="Alu"  
repeat\_region  
13107..13275  
/rpt\_family="B2"  
repeat\_region  
13518..13601  
/rpt\_family="B4"  
repeat\_region  
13532..13706  
/rpt\_family="B2"  
repeat\_region  
13743..13908  
/rpt\_family="L1"  
repeat\_region  
14060..14166  
/rpt\_family="L1"  
repeat\_region  
14150..15342  
/rpt\_family="L1"  
repeat\_region  
15729..16048  
/rpt\_family="MALR"  
repeat\_region  
17335..17449  
/rpt\_family="Alu"  
repeat\_region  
18836..19035  
/rpt\_family="L1"  
repeat\_region  
19041..19287  
/rpt\_family="B4"  
repeat\_region  
23075..23454  
/rpt\_family="MALR"  
repeat\_region  
23535..23935  
/rpt\_family="L1"  
repeat\_region  
24005..24090  
/rpt\_family="L1"  
repeat\_region  
24219..24279  
/rpt\_family="L1"

repeat\_region 25080..25179  
/rpt\_family="B4"  
repeat\_region 25485..25580  
/rpt\_family="L1"  
repeat\_region 25888..25959  
/rpt\_family="L1"  
repeat\_region 27683..27769  
/rpt\_family="L1"  
repeat\_region 27805..27957  
/rpt\_family="L1"  
repeat\_region 28112..28207  
/rpt\_family="Alu"  
repeat\_region 28720..29059  
/rpt\_family="L1"  
repeat\_region 29081..29175  
/rpt\_family="L1"  
repeat\_region 30059..30195  
/rpt\_family="Alu"  
repeat\_region 30255..30415  
/rpt\_family="L1"  
repeat\_region 30791..31062  
/rpt\_family="B2"  
repeat\_region 31275..31462  
/rpt\_family="B2"  
repeat\_region 31503..31701  
/rpt\_family="L1"  
repeat\_region 31751..31880  
/rpt\_family="L1"  
repeat\_region 32178..32531  
/rpt\_family="L1"  
repeat\_region 32667..32856  
/rpt\_family="L1"  
repeat\_region 32838..32933  
/rpt\_family="L1"  
repeat\_region 33219..33519  
/rpt\_family="L1"  
repeat\_region 33804..33873  
/rpt\_family="L1"  
repeat\_region 33875..34894  
/rpt\_family="L1"  
repeat\_region 34940..35077  
/rpt\_family="B2"  
repeat\_region 37436..37719  
/rpt\_family="B4"  
repeat\_region 38077..38225  
/rpt\_family="B4"  
repeat\_region 38295..38443  
/rpt\_family="MALR"  
repeat\_region 38655..38980  
/rpt\_family="MALR"  
repeat\_region 40754..40910  
/rpt\_family="L1"  
repeat\_region 41010..41141

Query Match 5.4%; Score 41.2; DB 9; length 198900;  
Best Local Similarity 49.5%; Pred. No. 0.56;  
Matches 106; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 48 TTTTCTCTATGCCGAGAGAGCTTTGTGTGCCGAATAGTTCGTAATATAGCACTTGGAATG 107  
DB 48619 TGAATTCACAGTGTGTAAAGGCTAATATCTGTAAAGTGTGACTTCAAGTAACAGAA 48560  
QY 108 CGAGGAGATTTTCCCTCGATTTTCACAATGATGTTGCGTAATATAGCACTTGGAATG 167  
DB 48559 TTACAGAAATGTTTCATGCTTTTATTTTAAATTAACCAAGACATTAATTGAGCATC 48500  
QY 168 GTTGTATAACTGACAGTAATATGACAGCGCGTCATTTCTGTCAGCTCAGACATGT 227  
DB 48499 CTAAAGGAATCTGCAATATGTCAGGACGAGGGAAGCATTTGAAATCTTAACAAATAA 48440  
QY 228 ACAAGTGACCTGGAGTTGCAACTTGGAAGAAGT 261  
DB 48439 ACTGTGTAGGTGTGTATAGTGTCTGGAATTAAT 48406



```

RESULT 10
AL450327/c
LOCUS DEFINITION
AL450327 174304 bp DNA linear PRI 18-MAY-2005
Human DNA sequence from clone RP11-173D14 on chromosome 6. Contains
part of a novel gene, a gene for a novel protein (likely ortholog
of mouse putative IKK regulator SIMPL), the 3' end of a gene for a
novel WD repeat-containing protein and a Cpg island, complete
sequence.

ACCESSION AL450327 AC027713
VERSION AL450327.12 GI:19572402
KEYWORDS HTG; Cpg island.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 174304)
Leonamornleart.D.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
On Mar 21, 2002 this sequence version replaced gi:16972990.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPeP; Info: Information
on the WormPeP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Ch6
RP11-173D14 is from the library RP11-1.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
Draft Sequence Produced by Whitehead Institute/MIT Center for
Genome Research, 320 Charles Street,
Cambridge, MA 02141, USA
http://www-seq.wi.mit.edu
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES
Source
1..174304
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-173D14"
/clone_1lb="RP11-1.1"
2000
/misc_feature
/note="Clone right end: RP1-263N12"
/locus_tag="RP11-173D14.3-001"
/locus_tag="RP11-173D14.3-001"
/complement(join(41639..44088,45046..45227))
/locus_tag="RP11-173D14.3-001"
/locus_tag="RP11-173D14.3-001"
/note="match: CDNA: AB056361 AB060929"
99396..130527
/locus_tag="RP11-173D14.1-001"
/join(95396..99815,117302..117367,129753..129883,
129988..130527)
/locus_tag="RP11-173D14.1-001"
/note="match: ESTs: AA068311 AA183360 AA270918 AT1239554
AT379905 AT478629 AT498886 AT917638 AK450696 BG614675
BB643845 BF056102 BF513853 BG604405 BG735351 BG773243
BB562425 BI667824 BI988366 BM127007 BM537735 BM808178
BM823374 BM830836 W62537
match: CDNA: AF093135 AJ242722 AK009948 AK013036 AK014712
AK020170 AK020448"
1299501..99815,117302..117367,129753..129883,
129988..130527)
/locus_tag="RP11-173D14.1-001"
/standard_name="OTTHUMP00000016770"
/note="match: proteins: Q9CRM0 Q9D615 Q9BSJ7 Q9UK12"
/codon_start=1
/protein_id="CAH70883.1"
/db_xref="GI:556525"
/db_xref="InterPro:IPR007497"
/translation="MSLQKTPPTPVPELVLPVADNRERNNLASGRTLLPGLRPLSST
QAOATPRVGVSTGSEVSAGPDRAOVVVSSTKRAAARAKSVCRLLDLYTQSLQOO
GVOAENITTKDPRVRYNAVMEAVCTFTFRKQNTCNFLVEKLDSSVVSPPQF
YHTPSVENLRQACLVAVENMRAQEVCLNVLGKPLIKESRTYEWKQIDDH
QSRSLSSLTVOOKIKSATIRHAAKVFIFFEYVGEKREKCHL"
120806..120926
/locus_tag="RP11-173D14.1-001"
/note="Single clone region. Assembly confirmed by
restriction digest data."
120970..120982
/locus_tag="RP11-173D14.1-001"
/note="Single clone region. Assembly confirmed by
restriction digest data."
121034
/locus_tag="RP11-173D14.1-001"
/note="Weak data"
130486..130491
/locus_tag="RP11-173D14.1-001"
130527
/locus_tag="RP11-173D14.1-001"
172305
/note="Clone left end: RP11-477B3"
/join(complement(ALJ56776..21:137647..137854),
complement(ALJ56776..21:137473..137531),
complement(ALJ56776..21:137318..137347),
complement(ALJ56776..21:137066..137125),
complement(ALJ56776..21:120286..120436),
complement(ALJ56776..21:120096..120194),
complement(ALJ56776..21:102461..102621),
complement(ALJ56776..21:85561..85782),
complement(ALJ56776..21:85137..85237),
complement(ALJ56776..21:78707..78717),
complement(ALJ56776..21:76964..77004),
complement(ALJ56776..21:76162..76260),
complement(ALJ56776..21:75248..75401),
complement(ALJ56776..21:74700..74824),
complement(ALJ56776..21:63348..63476),
complement(ALJ56776..21:61517..61742),
complement(ALJ56776..21:57872..58009),
complement(ALJ56776..21:57032..57215),
complement(ALJ56776..21:50486..50603),
complement(ALJ56776..21:47827..47967),
complement(ALJ56776..21:44970..45046),
complement(ALJ56776..21:42504..42735),
complement(ALJ56776..21:38210..38329),
complement(ALJ56776..21:30399..30506),
complement(ALJ56776..21:29667..29761),
complement(ALJ56776..21:29453..29535),
complement(ALJ56776..21:28563..28674),
complement(ALJ56776..21:25321..25382),
complement(ALJ56776..21:22715..22870),
complement(ALJ56776..21:21308..21448),
complement(ALJ56776..21:18093..18218),

```

## gene

## mRNA

```
complement (AL356776.21:15231..15300),
complement (AL356776.21:14850..14900),
complement (AL356776.21:14432..14501),
complement (AL356776.21:7241..7393),
complement (AL356776.21:6329..6492),
complement (AL356776.21:5619..5878),
complement (AL356776.21:4918..5115),
complement (172470..173254))
/locus_tag="RP11-173D14.2-001"
join (complement (AL356776.21:42504..42561),
complement (AL356776.21:38210..38329),
complement (AL356776.21:30399..30506),
complement (AL356776.21:29667..29721),
complement (AL356776.21:29453..29535),
complement (AL356776.21:25563..25674),
complement (AL356776.21:25321..25382),
complement (AL356776.21:22715..22870),
complement (AL356776.21:15231..18218),
complement (AL356776.21:14850..14900),
complement (AL356776.21:14432..14501),
complement (AL356776.21:7241..7393),
complement (AL356776.21:6329..6492),
complement (AL356776.21:5619..5878),
complement (AL356776.21:4918..5115),
complement (172470..173254))
/locus_tag="RP11-173D14.2-002"
join (complement (AL356776.21:137647..137854),
complement (AL356776.21:137473..137531),
complement (AL356776.21:137318..137347),
complement (AL356776.21:137065..137125),
complement (AL356776.21:120096..120194),
complement (AL356776.21:120286..120436),
complement (AL356776.21:120461..108221),
complement (AL356776.21:85561..85782),
complement (AL356776.21:85137..85237),
complement (AL356776.21:78707..77201),
complement (AL356776.21:77101..77004),
complement (AL356776.21:76964..76260),
complement (AL356776.21:75248..75401),
complement (AL356776.21:74700..74834),
complement (AL356776.21:63348..61742),
complement (AL356776.21:61517..58009),
complement (AL356776.21:57872..57215),
complement (AL356776.21:50486..50603),
complement (AL356776.21:47827..47967),
complement (AL356776.21:44970..45046),
complement (AL356776.21:42504..42735),
complement (AL356776.21:38210..38329),
complement (AL356776.21:30399..30506),
complement (AL356776.21:29667..29721),
complement (AL356776.21:29453..29535),
complement (AL356776.21:25563..25674),
complement (AL356776.21:25321..25382),
complement (AL356776.21:22715..22870),
complement (AL356776.21:21308..21428),
complement (AL356776.21:21308..21428),
```

Query Match Best Local Similarity 5.3%; Score 40.2; DB 8; Length 174304;

Matches 60; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
QY 347 AAAGACAGTGAAGAAGACAGTTCTGCAACAATATCTGTTGGTGGTGAAGCGCATCG 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11062 AAAGACAGAAAGTAGAGAGTGCTCTAACAACAATCCAAAGATTGATGATGTTGATCA 11003
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 407 CTGGGATTTATGAAGAAGCAGAACTTGACCA 439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11002 CTGGATTTATATATATGCTTACATTTGCTAAA 10970
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 11

```
166494/c
LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
VERSION 166494.1 GI:2724471
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 7218)
  Unclassified.
AUTHORS
  Donner, F., Schefflinger, F. and Falkner, F. Gunter.
TITLE
  Recombinant fowlpox virus
JOURNAL
  Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
  source
    location/Qualifiers
    1..7218
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
```

Query Match Best Local Similarity 5.1%; Score 39.2; DB 6; Length 7218;

Matches 32; Conservative 197; Mismatches 185; Indels 0; Gaps 0;

```
QY 102 ACAGCGCAGCAGCATTTTCTCGATTTCACATGATGTCCTCGAATATAGACCTTG 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1505 AAACGCGATGAGCATCAGTGAATTAATCTATCTATGCAAGTAAAGATAGAA 1446
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 162 GAATGTTTGAATTAACGACAGTAATGACAGCGGTCTTGTGTCAGCTCAGA 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1445 GAATTTGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1386
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 222 ACATGTAACAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1385 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 282 CATGCAATGATCCCTACCGATTAATACAGCTGCTCAAAATTCGTAGATGGCTGT 341
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1345 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 342 ACTTAAAGACAGTGAAGAAGACAGTTCGCAACAAGTATCTGTAATGCGTAGCG 401
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1265 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 402 CATCGCTGAGATTTTGAAGAAGCAGCAAACTTGACCAATTTGCTTACAGAGTGGCTGAAC 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1205 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1146
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 462 CTTCAGCAAACTTGCAACTATAGAAACCGAAAGTTGATGTCCTCATTAGA 515
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1145 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 12

AC022202 AC022202 95097 bp DNA linear PRI 08-OCT-2003

LOCUS Homo sapiens BAC clone RP11-321B8 from 7, complete sequence.

DEFINITION AC022202

ACCESSION AC022202

VERSION AC022202.12 GI:13992747

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 95097)

AUTHORS Sulem, J.B. and Wilson, R.

TITLE Toward a complete human genome sequence

JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

PUBMED 9847074

REFERENCE 2 (bases 1 to 95097)

AUTHORS Abbott, A., Hawkins, M. and Meyer, R.

TITLE The sequence of Homo sapiens BAC clone RP11-321B8

JOURNAL Unpublished (2001)

REFERENCE 3 (bases 1 to 95097)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JAN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 4 (bases 1 to 95097)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 95097)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 6 (bases 1 to 95097)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 7 (bases 1 to 95097)  
 AUTHORS Wilson, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (08-OCT-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On May 8, 2001 this sequence version replaced gi:13431144.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu>  
 Contact: [saplens@wustl.wustl.edu](mailto:saplens@wustl.wustl.edu)  
 Summary Statistics  
 Center project name: H\_NH0321E08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, B., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pSPACE3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-468B6, 200 bp overlap the clone sequenced to the right is RP11-561N12. Actual start of this

FEATURES	clone is at base position 61428 of RP11-468B6 actual end is at base position 95097 of RP11-321E8.
source	Location/Qualifiers
	1. 95097
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="RP11-321E8"
	/clone_11b="RPCT-11"
repeat_region	273..564
	/rpt_family="Alu"
repeat_region	1280..1352
	/rpt_family="L2"
repeat_region	1371..1404
	/rpt_family="MER1_type"
repeat_region	1570..1808
	/rpt_family="L2"
repeat_region	1861..2172
	/rpt_family="Alu"
repeat_region	3218..3294
	/rpt_family="CR1"
repeat_region	3320..3610
	/rpt_family="Alu"
repeat_region	4987..5293
	/rpt_family="Alu"
repeat_region	5127..6021
	/rpt_family="Alu"
repeat_region	6134..6454
	/rpt_family="Alu"
repeat_region	6798..6945
	/rpt_family="MER1"
repeat_region	9181..9201
	/rpt_family="AT-rich"
repeat_region	9242..9471
	/rpt_family="L1"
repeat_region	9472..9785
	/rpt_family="Alu"
repeat_region	9786..9898
	/rpt_family="L1"
repeat_region	9907..10485
	/rpt_family="ERV1"
repeat_region	10594..10760
	/rpt_family="L1"
repeat_region	11288..11311
	/rpt_family="TMAA"
repeat_region	13027..13325
	/rpt_family="Alu"
repeat_region	14024..14329
	/rpt_family="Alu"
repeat_region	14395..14416
	/rpt_family="CMA"
repeat_region	14818..14837
	/rpt_family="GAAA"
repeat_region	15486..15770
	/rpt_family="Alu"
repeat_region	16662..16946
	/rpt_family="L1"
repeat_region	18463..18615
	/rpt_family="MER1_type"
repeat_region	18605..18650
	/rpt_family="MER1"
repeat_region	18773..18987
	/rpt_family="L1"
repeat_region	19997..20162
	/rpt_family="MER1_type"
repeat_region	20282..20582
	/rpt_family="Alu"
repeat_region	20586..20743
	/rpt_family="L1"
repeat_region	20770..21530
	/rpt_family="L1"

```

repeat_region 21732..21798
                /rpt_family="L1"
repeat_region 21799..22104
                /rpt_family="Alu"
repeat_region 22105..22878
                /rpt_family="L1"
repeat_region 22879..22902
                /rpt_family="AT-rich"
repeat_region 22934..22957
                /rpt_family="AT-rich"
repeat_region 23000..23370
                /rpt_family="Alu"
repeat_region 23481..23523
                /rpt_family="AT-rich"
repeat_region 23738..23769
                /rpt_family="AT-rich"
repeat_region 24080..24144
                /rpt_family="L1"
repeat_region 24145..24452
                /rpt_family="Alu"
repeat_region 24453..25442
                /rpt_family="L1"
repeat_region 26505..26544
                /rpt_family="TTTG)n"
repeat_region 27297..27572
                /rpt_family="L1"
repeat_region 27573..27871
                /rpt_family="Alu"
repeat_region 27872..28013
                /rpt_family="L1"
repeat_region 28014..28325
                /rpt_family="Alu"
repeat_region 28326..28773
                /rpt_family="L1"
repeat_region 28777..28812

```

Query Match 5.1%; Score 39; DB 8; Length 95097;  
 Best Local Similarity 53.6%; Pred. No. 2.6; Indels 0; Gaps 0;  
 Matches 81; Conservative 0; Mismatches 0;

```

QY 82 GAATGCTTAAGCAAGTACAGCGAGGAGATTTCTGATTTTCAACATGATGT 141
DB 11138 GCAAACTTTATGACCTTTCCAAAGAACACATATTTTCCAGTATTTCAAAAAAAT 11197
QY 142 TCGTCGAAATATAGCACTTGGAATGTTGATTAATCGACATGTAATGCAGCGGT 201
DB 11198 TCAATGAAATGTCAGCTTTAAAGACTTTTAAATAATAGTTTAAAGAGCTCAAGA 11257
QY 202 CATTCTTGCTCCAGCTCAGAACATGTCAAA 232
DB 11258 GATGCAAGAGAAATCTGAAAACCAATATCAAA 11288

```

RESULT 13  
 AC121380\_2/c  
 WPCOMMENT  
 Sequence split into 4 fragments LOCUS AC121380 Accession AC121380

Fragment Name	Begin	End
AC121380_0	1	110000
AC121380_1	100001	210000
AC121380_2	200001	310000
AC121380_3	300001	407488

Continuation (3 of 4) of AC121380 from base 200001 (AC121380 Rattus norvegicus clone CH2

Query Match 5.1%; Score 38.6; DB 14; Length 110000;

Best Local Similarity 47.7%; Pred. No. 3.5; Indels 124; Indels 0; Gaps 0;

```

QY 126 ATTTTCATGATGTTGCTGCAATATAGCACTTGGAATGTTGATTAAGTGAAG 185
DB 30506 AGTATCAAAAGCAATTTAAGTATATAGCATATATAAATAATTAATGATTA 30447
QY 186 TAAATGACAGCGGCTATTTGCTCAGCTCAGAACATGTCACAAAGTGAAGTGGATT 245

```

```

DB 30446 TAAATTTATGAGACTTTTGTGACTTAGCTAAATTTTCAAAATTTTCAAAAT 30387
QY 246 GCAACTGGAAGAGTGCAGCAACAGATTTGGCAGATGATGATCCCTACCGATA 305
DB 30386 ATAAATTTTACCAATTAATTAATTTGTAAGAAATGAAATTAATTTTCACTTAAGAAATT 30327
QY 306 ATACCAGCTGCTGCAAAATATGCTGATAGTGTGATCTTCAAGACATGAGAG 362
DB 30326 TCAATTCAGAGGGAATTTGAATGTCATGCTTAACAGAAATTCACATTAATCAG 30270

```

RESULT 14  
 AC120958/c  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-25B23, \*\*\* SEQUENCING IN PROGRESS

AC120958 225475 bp DNA linear HTG 19-NOV-2002  
 HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.  
 SOURCE  
 ORGANISM  
 Rattus norvegicus (Norway rat)

# REFERENCE AUTHORS

Allen, C., Allen, H., Albrooke, S., Anlin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnes, M., Behan, F., Biewald, K., Blair, J., Blankensbury, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, D., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., d'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Derramo, C., Ding, Y., Dinh, R., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Garcia, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulky, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loubeed, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manthey, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Mavrosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okunnu, G., Olarunbasogun, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, A., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Fu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodery, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shecty, J., Shvartsbeyn, A., Slisom, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svetek, A., Tabot, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinsley, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, D., Yoon, L., Yoon, V., Yu, F., Zhang, D., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhauser, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstein, G., and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 225475)  
 Worley, K.C.  
 Direct Submission  
 Submitted (14-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 225475)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 COMMENT  
 On Nov 19, 2002 this sequence version replaced gi:22856233.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GWVW  
 Center clone name: CH230-25823  
 ----- Summary Statistics  
 Assembly program: Phrap, version 0.990329  
 Consensus quality: 206303 bases at least Q40  
 Consensus quality: 208669 bases at least Q30  
 Consensus quality: 210148 bases at least Q20  
 Estimated insert size: 212963; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 1 225475: contig of 225475 bp in length.  
 FEATURES  
 source  
 1. 225475  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-25823"  
 1. 2285  
 misc\_feature  
 /note="wgs contig"  
 misc\_feature  
 /note="wgs contig"  
 /note="wgs contig"  
 /note="wgs contig"  
 ORIGIN  
 Query Match 5.1%; Score 38.6; DB 14; Length 225475;  
 Best Local Similarity 47.7%; Pred. No. 3.6;  
 Matches 113; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 26 GTATGTCGTCCTCTCCGCGCCGCTGTCCTTATGCGAAGCAGGCTTTGTCGGAAT 85  
 |||||  
 Db 92264 GGAGCAGGGGGCTCATTTTCCCTGTCGATCTCGGAGTTGACAGGTGTGACGAC 92205  
 QY 86 AGCTTAAGCCAAAGTGAAGCGCGAGGAGATTTTCTCGATTTTCAATGATGTTGCT 145  
 |||||  
 Db 92204 TCGAGTAAGACATTTTACCTGCTGAGCCACATGCTGCGCCCTTATTCCTGTTTGTCTT 92145  
 QY 146 CGAAATATGACACTGGAATATGTTGATTAATCTGACAGTAATACCAACCGGTCATT 205  
 |||||  
 Db 92144 TGAATAATTTTTTTTAAATTAAGTTTAACTTATTAAGTAATCACTATTTCTTGATA 92085  
 QY 206 CTGTGTCAGCTCAGAACATGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 262  
 |||||  
 Db 92084 CTGTGAGCAACTGAGATGATGATTAATCTTAATCTGATTTGATCAATCAATCA 92028  
 RESULT 15  
 AC121042  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-415F5, WORKING DRAFT SEQUENCE, 3  
 unnumbered pieces.  
 AC121042  
 ACCESSION  
 AC121042.4 GI:24635446  
 VERSION  
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.  
 KEYWORDS  
 Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.  
 1 (bases 1 to 239001)  
 REFERENCE  
 AUTHORS  
 Muzny, D., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Ayalew, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaru, D., Bandaru, D., Barber, M., Barnstead, M.,  
 Bismuth, K., Blair, J., Blankenship, K., Blyth, P., Brown, M.,  
 Bryant, N., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Ande, C., Dederich, D.,  
 Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M.,  
 Hollins, B., Howells, S., Hui, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowitz, C., Kraft, C.L., Lebow, H., Lewis, L., Lewis, L., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenshuber, L., Louisedge, H., Lozano, R.J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindaratne, M., Mamoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, K., Martin, E.,  
 Manning, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwacke, O., Okwou, G., Olamugbo, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plapper, F., Polindexter, A., Popovic, D., Primus, B., Pu, L.,  
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Slisom, I., Sitter, C.D., Smajic, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svetek, A., Taber, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K.,

Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Morley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, L., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Unpublished  
Direct Submission  
2 (bases 1 to 239001)  
Morley, K. C.

Submitted (15-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 239001)  
Rat Genome Sequencing Consortium.

Direct Submission  
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 6, 2002 this sequence version replaced gi:23908010.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: CH230-415FS  
Center clone name: GWOM  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 172958 bases at least Q40  
Consensus quality: 175322 bases at least Q20  
Consensus quality: 176115 bases at least Q20  
Estimated insert size: 174254; sum-of-contigs estimation  
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 236295: contig of 236295 bp in length  
\* 236296 236395: gap of unknown length  
\* 236396 237724: contig of 1329 bp in length  
\* 237725 237824: gap of unknown length  
\* 237825 239001: contig of 1177 bp in length.  
Location/Qualifiers  
1. 239001  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-415FS"  
1. 1011  
misc\_feature  
/note="wgs\_end\_extension"

misc\_feature  
48358..117554  
/note="clone\_boundary  
clone\_end:T7  
site:  
end sequence: BZ193117"  
complement (179558..180307)  
/note="clone boundary  
clone\_end:Sp6  
site:  
end sequence: BZ193118"  
181460..182492  
/note="wgs\_end\_extension  
clone\_end:Sp6"  
235097..236295  
/note="wgs\_end\_extension  
clone\_end:Sp6"  
236296..236395  
/estimated\_length=unknown  
237725..237824  
/estimated\_length=unknown

ORIGIN  
Query Match 5.1%; Score 38.6; DB 14; Length 239001;  
Best Local Similarity 47.7%; Pred. No. 3.6;  
Matches 113; Conservative 0; Mismatches 124; Indels 0; Gaps 0;  
QY 126 ATTTTCAGATGATGCTTCGTGAATATGACACTTGGAAATGTTGATTAATGACAG 185  
DB 55869 AGTATCAAAAGCATTATTAAGTATATGATATGATTAATTAATTAATTAATGATTA 55928  
QY 186 TAAATGACAGCGCGTCATTTCTTGTCACGCTCAGAACATGTACAAAGTGAATGAGATT 245  
DB 55929 TAAATTTATTTATGGAACCTTTTGTGTGACCTTAGCTAAATTTACAAATTTACCAAATT 55988  
QY 246 GCAACTTGAAAGAGTACAGACACACAGATTGGCCATGCAATGATCCCTACCGATTA 305  
DB 55989 ATTAATTTTACCAATTTACCAATTTGTAAGAAATGAAATTAATTAATTAATTAATTAATTA 56048  
QY 306 ATACAGCTTCGCTCAAAATATTCGCTAGATGCGTGTACTTCAAGACAGTGAAGAG 362  
DB 56049 TCAGTATCAAGGGGAAATTTGAATGTCTAGCTTAACAGAAATTCCTAATATG 56105

Search completed: March 4, 2006, 10:13:53  
JOB time : 4218 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:02:49 ; Search time 589 Seconds  
(without alignments)  
8633.549 Million cell updates/sec

Title: US-10-527-771-9

Perfect score: 763  
Sequence: 1 gagaaactgcatgctgcgcgcg.....tcacgcaaaaaaaaaaaaaa 763

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001as:\*
- 5: Geneseq2001bs:\*
- 6: Geneseq2002as:\*
- 7: Geneseq2002bs:\*
- 8: Geneseq2003as:\*
- 9: Geneseq2003bs:\*
- 10: Geneseq2003cs:\*
- 11: Geneseq2003ds:\*
- 12: Geneseq2004as:\*
- 13: Geneseq2004bs:\*
- 14: Geneseq2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	763	100.0	763	12	ADK13846
2	273.4	35.8	828	12	ADK13844
3	41.4	5.4	2000	8	ADA71938
4	37.4	4.9	2000	8	ADA71938
5	36	4.7	260803	13	ABD32730
6	35.4	4.6	1537	6	AB897663
7	35.4	4.6	7769	6	AB996867
8	34.6	4.5	110000	13	ABD32627_0
9	34.6	4.5	194883	11	ACN44398
10	34.4	4.5	1003	6	ABQ24108
11	34.4	4.5	1003	6	ABQ24109
12	34.2	4.5	110000	13	ABD32806_1
13	34	4.5	3457	4	AB110900
14	34	4.5	4092	4	AB118190
15	34	4.5	7160	4	AB113492
16	33.6	4.4	2000	6	AB216370
17	33.6	4.4	130263	6	ABK83573
18	33.4	4.4	1321	14	ADY37283
19	33.4	4.4	11204	4	AA104557

## ALIGNMENTS

20	33.4	4.4	11204	4	ABL97480	Ab197480 Human tes
21	33.4	4.4	19802	4	AAK90894	AAK90894 Human dig
22	33.4	4.4	19802	4	AA104554	AA104554 Human rep
23	33.4	4.4	19802	4	ABL97477	ABL97477 Human tes
24	33.4	4.4	21403	4	AAK90893	AAK90893 Human dig
25	33.4	4.4	21403	4	AA104555	AA104555 Human rep
26	33.4	4.4	21403	4	ABL97478	ABL97478 Human tes
27	33.4	4.4	23763	4	AAK90895	AAK90895 Human dig
28	33.4	4.4	23763	4	AA104558	AA104558 Human rep
29	33.4	4.4	23763	4	ABL97481	ABL97481 Human tes
30	33.4	4.4	67403	13	ADS36460	ADS36460 Human aut
31	33.4	4.4	97662	4	AAK83908	AAK83908 Genomic s
32	33.4	4.4	106746	3	AA10225	AA10225 Human PCT
33	33.2	4.4	530	8	ABX98607	ABX98607 Rice leaf
34	33	4.3	804	12	ADN72288	ADN72288 Thale cre
35	33	4.3	1193	13	ADK31759	ADK31759 Plant ful
36	33	4.3	1510	3	AAK42522	AAK42522 Arabidops
37	33	4.3	1884	11	ADN39685	ADN39685 Cancer/an
38	33	4.3	2876	11	ADN39684	ADN39684 Cancer/an
39	33	4.3	44978	13	ADS36458	ADS36458 Human aut
40	32.8	4.3	3100	6	ADK22128	ADK22128 Arabidops
41	32.6	4.3	796	10	ADK58398	ADK58398 Plant DNA
42	32.6	4.3	796	10	ADK54354	ADK54354 Plant DNA
43	32.6	4.3	801	10	ADK57570	ADK57570 Plant DNA
44	32.6	4.3	1278	6	ADK33756	ADK33756 Mouse Mtg
45	32.6	4.3	1278	12	ADH08544	ADH08544 DNA seque

## RESULT 1

ID ADK13846 standard; cDNA: 763 BP.

ADK13846;

03-JUN-2004 (first entry)

Ostertagia ostertagi immunogenic protein cDNA seq id 9.

antiparasitic; gene therapy; vaccine; nematode; worm;

Ostertagia ostertagi protein; immunogenic protein; gene; ss.

Ostertagia ostertagi.

Key Location/Qualifiers

CDS 11.706 /tag= a /product= "Ostertagia ostertagi immunogenic protein"

US2004052817-A1.

18-MAR-2004.

13-SEP-2002; 2002US-00243319.

13-SEP-2002; 2002US-00243319.

(GELD/) GELDHOF P.

(VERC/) VERCAUTEREN I.

(DMAE/) DE MAERE V.

(CLAER/) CLAERBOUW E.

(VERC/) VERCRUYSE J.

Geldhof P, Vercauteren I, De Maere V, Claerebout E, Vercruyse J;

WPI: 2004-247704/23.

P-PSDB; ADK13847.

New Ostertagia ostertagi proteins and nucleic acid sequences, useful for diagnosing, preventing or treating infections caused by Ostertagia ostertagi in animals.





Db 164 GGTCAGAGCCGTTGCTCAACTCACCGAGCTGTTCAATGCGAAATGTTCTCGGTCCA 2232

Qy 215 GCTCAGAACATGTACAAAGTGA CTGGGAATTGCACTTGGAGAGTAGACACAAG 2747

Db 224 GCTAAGAACATGTACAGAAATGCACTGGGACTGGCAATCGAAGCAAAAGCAAGCGAATG 2833

Qy 275 ATTGGCCCATGCAATATCCCTTAACCGATTAATACAGCCTGCTCAAAATATCGTTAGA 3344

Db 284 ATTTGGCCCATGCACTAGCCTCTGCCAAATAGACAGAGTATTCACAAATCTCGCTCAR 3433

Qy 335 TGGCTGTACTTCAAAGACAGTGAAGAGACAGTTCTGCACAAGTATGTGGTATTGG 3944

Db 344 TGGCTATCTTTCCAAACAGTCAGAAAMTGAAGTGTGACGAAACGCCCTGCTTTGG 4030

Qy 395 GTGAGCGCATCGCTGGGAATTTATGAAGGACAGAACTTGACCAATTTGCTAACCGGTGG 4544

Db 404 GTTACCGCATCACTACGAATCTTCAACTGATACAGAACTTAACATTTAACTGGCAA 4633

Qy 455 GCTGAACCTCTAGCAAAACATTTGCAAACTATAGAAACGAAAGGTGGATGTGCCATPAG 5144

Db 464 ATTAGACCATATCCAACTTTGGGAATCGCAAAACCTTAATAATTGGATGTGCTCACAAA 5233

Qy 515 ATCTG-----CCCGCGCTCAGCAAAACATGTAATCTGTCGTTATGAGACCCCAA 5688

Db 524 GTGTGAATATTCGCCACCGGGAATAATATGTTGTGTCTTGGCGCTTAATGGCGGCGAAGTA 5833

Qy 569 CTTCGACCGAAGCAAGTTATCTGGCAGAAAGAAAGGCTTGTGTGTCGACGCTCGTCCA 6288

Db 584 CTCCAAGATTAAGAAAGTTGATATGGGACAAAGGGACCAACTTGCAATGTCATATGCTATCCC 6433

Qy 639 GATTCAATCTTGTCGACAACAACCTGTGTGACAGCGGAGATGTCTGCCAGTGTGCCACAG 6688

Db 644 AACTGGTTCTGTGCAACAATCTGTGTGACACATATGCTGCTCGACACATTCGCCAAGCAG 7033

Qy 689 TGTTCGCGGTGGCATGAGCGCAAAAGAAATTGTAATGCAACC 731

Db 704 CCTTGTAAATTCACCTTGAAGCAAAAGCGGTGTGTATGTCCC 746

PT	Identifying at least one gene involved in plant resistance or response to bacterial infection for conferring resistance or tolerance to a plant to pathogenic, fungal or viral infection by determining or detecting plant
DR	WPI; 2003-175290/17.
XX	
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.
XX	
PL	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
P1	Katesgiri F, Quan S, Tao Y, Whitam S, Xie Z, Zhu T, Zou G;
XX	
XX	
OS	Oryza sativa.
XX	
PN	WO200300898-A1.
FD	03-JAN-2003.
XX	
PF	22-JUN-2001; 2001WO-IB001105.
XX	
PR	22-JUN-2001; 2001WO-IB001105.
XX	
DE	Rice gene, SEQ ID 5263.
DT	20-NOV-2003 (first entry)
AC	ADA71938;
XX	
ID	ADA71938 standard; DNA; 2000 BP.
RESULT 3	

PT gene expression.  
 XX  
 XX Claim 27; SEQ ID NO 5263; 89pp; English.  
 XX  
 CC The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 CC

Seq	Sequence	2000 BP	336 A	265 C	284 G	363 T	0 U	752 Other									
Query	Match	5.4%	Score 41.4	DB 8	Length 2000												
Best	Local Similarity	9.3%	Pred. No. 0.066														
Matches	54	Conservative	263	Mismatches	264	Indels	2	Gaps 1									
Qy	167	GGTTGA	AAACTG	GCAG	TAAATG	CAAC	AGCG	GTCA	TCTT	GTGTC	AGCT	CA	GTC	GA	CA	TG	226
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	219	RKGYKKM	YTYSA	SRCB	AWTTS	YSMA	CSYTM	CRSR	BRB	MMKMR	KMR	KMS	NSY	GY			278
Qy	227	TACAA	TGTC	AGTC	GGAT	TTC	CA	CTTG	CA	TTG	GA	AGA	AGT	GA	CA	CA	286
		:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	:: ::	
Db	279	SWSYKMM	CTAYK	KSYS	RYMC	YRG	GGWR	GTAT	RTW	GGYMS	RRMA	MMY	KDM	MY	RG	KCM	338
Qy	287	AATG	ATCC	CCAC	CA	GTAA	TATAC	CA	CGCT	GGT	CAAA	TATG	CT	AG	TAG	AGCT	346
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	339	KRGW	AGMM	RRSR	CMCR	SKACTY	RRRR	RRRR	RR	RR	RR	RR	RR	RR	RR	RR	398
Qy	347	AAAG	CA	GTGA	GA	GAG	AC	AGT	TTC	GA	CA	AG	TAT	CGT	GTAT	GGT	406
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	399	RMRG	YSR	MR	SR	CKRA	RM	MR	CR	SG	AA	KK	CG	CM	T	CR	458
Qy	407	CTG	GA	TTT	AA	GA	AG	CA	AA	CTT	GA	CCA	ATT	GT	CTA	AA	466
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	459	RMYR	MR	KK	CG	SR	TTM	GGT	RR	GA	MM	GT	MR	CR	YK	RR	518
Qy	467	GCAA	CA	ATTG	CA	AACT	AT	AG	AA	CCG	AA	AG	GTG	G	-	ATG	524
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	519	SAR	TYM	Y	CA	KK	YS	AA	RR	AR	RC	MY	RR	K	Y	MM	578
Qy	525	CTCA	G	CA	AA	CA	TG	TA	GT	AT	CT	CG	T	GT	A	TG	584
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	579	SWY	CK	MS	TY	AS	CM	S	AR	KA	GA	CM	CR	S	MS	AS	638
Qy	585	TTAT	CTG	CA	GGA	GA	GA	AG	CG	CTT	GTG	TG	CG	CA	CG	CT	644
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	639	MTSS	RR	SR	MR	K	SY	TC	MY	RR	K	MS	K	MS	T	CT	698
Qy	645	ACA	AC	CTG	TG	TA	CA	CA	CG	GA	GT	CTG	CA	GTG	CC	CA	704
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	699	YTR	YS	Y	MT	Y	AA	Y	T	S	T	RR	AA	T	GM	K	758
Qy	705	GA	AG	CG	AAA	GA	AA	TTG	T	AG	T	CA	C	CC	CG	CA	747
		:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	:: :: ::	
Db	759	KMR	RR	Y	AT	RR	MM	MM	Y	R	S	M	K	Y	T	C	801

XX	RESULT 4
XX	ADA71938/c
XX	ID ADA71938 standard; DNA; 2000 BP.
XX	AC ADA71938;
XX	DT 20-NOV-2003 (first entry)
XX	DE Rice gene, SEQ ID 5263.
XX	LN Plant; bacterial infection; fungal infection; viral infection; rice,
XX	TM



Query Match 4.7%; Score 36; DB 13; Length 260803;  
 Best Local Similarity 45.7%; Pred. No. 38;  
 Matches 126; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 103 CAGCCGAGGAGGATTTCTCGATTTCACATGATGTCGTAATATAGCACTTG3 162  
 DB 259473 CTGGGAAAGATATCTATCCGCCCTTATCCTGAAATGATGACGAATTTGGCACTAT 259532  
 QY 163 AATGTTTGAATTAACGTAGACAGTAAATGAGAGCGGGTCACTTGTGTCAGCTAGAA 222  
 DB 259533 GATGTAATGGTATGTCAGCTTAAATGCTTTCATGCTTATCATGTGTTAAGCACTA 259592  
 QY 223 CATGTCAAAAGTGAAGCTGGGATTTGCAATTTGAAAGATAGCAGCAACAGATTCGCC 282  
 DB 259593 AATATTGGGGGTGCTTTGTTTACGACGATGACAGATTAACCTGACTGATTCCTAAC 259652  
 QY 283 ATGCAATGATTCCTTACCAATTAATACAGCTGGCTCAAAATATGCTAGATGCTGTA 342  
 DB 259653 ATGCACTCAGCATTTTGAACAAATCCAGCTATCAGTATTTCACTTTCCTTTGATTGTA 259712  
 QY 343 CTTCAAGACAGTGAAGAGACATTTCTGCANCA 378  
 DB 259713 CCAGATCCATTAAAGAGGACAGAAATTCACANCA 259748

RESULT 6  
 ABS97663/c  
 ID ABS97663 standard; DNA; 1537 BP.  
 AC ABS97663;  
 XX  
 XX 23-DEC-2002 (first entry)  
 DE Human glutathione-S-transferase 12 (GST12) gene sequence.  
 XX Human; db; cytochrome P450 A1; CYP4501A1; UGT2B4; MDR1;  
 XX cytochrome P450 A2; CYP4501A2; cytochrome P450 02B; CYP45002B1; LTP;  
 XX adrennergic receptor beta1; ADRB1; aryl hydrocarbon; AHR; MRP3; NR112;  
 XX aryl hydrocarbon receptor nuclear translocator; ARNT; catepsin S; CTSS;  
 XX cyclooxgenase 2; COX2; diazepam binding inhibitor; DBI; haematological;  
 XX epoxide hydroxylase 2; EPHX2; 5-lipoxygenase activating protein; FLAP;  
 XX glutathione-S-transferase 12; GST12; histamine-N-methyl transferase;  
 XX HMMT; kallikrein 2; KLK2; nicotinamide-N-methyl transferase; NNMT;  
 XX NADPH quinone oxidoreductase 2; NQO2; sulfoxtransferase thermolabile; STM;  
 XX UDP-glucuronosyl transferase 2B4; UDP-glucuronosyl transferase 2B7;  
 XX UGT2B7; UDP-glucuronosyl transferase; UGT2B15; urokinase receptor; UPA;  
 XX multidrug resistance 1; lactotransferrin; orphan nuclear receptor;  
 XX acetylcholine muscarinic receptor; CHMR1; CHMR2; CHMR3; CHMR4; CHMR5;  
 XX altered drug metabolism; cardiovascular function; colorectal tumour;  
 XX central nervous system; pulmonary; immunological; SNP;  
 XX single nucleotide polymorphism.  
 XX Homo sapiens.  
 XX OS  
 XX MO200257410-A2.  
 XX PD 25-JUL-2002.  
 XX PF 28-NOV-2001; 2001WO-US044838.  
 XX PR 28-NOV-2000; 2000US-00724389.  
 XX (DNAS-) DNA SCI LAB INC.  
 XX Guida M, Hall J;  
 XX MPI; 2002-69852/75.  
 PT Isolated nucleic acid molecules having polymorphisms in known human genes  
 PT e.g. cytochrome p450 and catepsin S useful as genetic linkage markers  
 PT for locating, identifying and characterizing the genes responsible for

PT disorder-related traits.  
 XX  
 PS Example 12; Page 321; 714pp; English.  
 XX  
 CC This invention relates to the sequence of an isolated nucleic acid  
 CC molecule comprising at least one base variation from that of a known  
 CC human cytochrome P450 A1 (CYP4501A1), cytochrome P450 A2 (CYP4501A2),  
 CC cytochrome P450 02B1 (CYP45002B1), adrennergic receptor beta1 (ADRB1),  
 CC aryl hydrocarbon (AHR), aryl hydrocarbon receptor nuclear translocator  
 CC (ARNT), catepsin S (CTSS), cyclooxgenase 2 (COX2), diazepam binding  
 CC inhibitor (DBI), epoxide hydroxylase 2 (EPHX2), 5-lipoxygenase activating  
 CC protein (FLAP), glutathione-S-transferase 12 (GST12), histamine-N-methyl  
 CC transferase (HMMT), kallikrein 2 (KLK2, nicotinamide -N-methyl  
 CC transferase (NNMT), NADPH quinone oxidoreductase 2 (NQO2),  
 CC sulfoxtransferase thermolabile (STM), UDP-glucuronosyl transferase 2B4  
 CC (UGT2B4), UDP-glucuronosyl transferase 2B7 (UGT2B7), UDP-glucuronosyl  
 CC transferase (UGT2B15), urokinase receptor (UPA), multidrug resistance 1  
 CC (MDR1), lactotransferrin (LTF), multidrug resistance associated protein 3  
 CC (MRP3), orphan nuclear receptor (NR112), or acetylcholine muscarinic  
 CC receptor 1, 2, 3, 4, or 5 (CHMR1, CHMR2, CHMR3, CHMR4 or CHMR5) sequence.  
 CC The polymorphisms in the human genes cited in the invention are useful as  
 CC genetic linkage markers for locating and characterizing the genes that  
 CC are responsible for specific traits within the genome and eventually  
 CC identifying the genes responsible for a variety of disorder-related  
 CC traits as a result of their e.g., overexpression, constitutive  
 CC expression, mutation or underexpression, which may be used in diagnosing  
 CC and/or treating the disorders. The nucleic acid molecules comprising the  
 CC polymorphic sequences contained in CYP4501A1, CYP4501A2, CYP4502B1,  
 CC ARNT, EPHX2, GST12, NNMT, NQO2, NR112, STM, UGT2B4, UGT2B7, UGT2B15, AHR,  
 CC MDR1 and/or MDR3 are useful for screening individuals for altered drug  
 CC metabolism. The polymorphic sequences contained in CYP4501A1, CYP4501A2,  
 CC AHR, MDR1 and/or MDR3 may also be used to screen individuals for  
 CC susceptibility to cancer. Polymorphic sequences in ADRB1 or CHMR2 are  
 CC used to screen for altered cardiovascular function. In COX2 for altered  
 CC susceptibility to colorectal tumours, in DBI or CHMR1 for altered central  
 CC nervous system function, in FLAP and HMMT for altered pulmonary,  
 CC immunological or haematological function, in KLK2 for altered seive  
 CC protease activity in the prostate, in LTP for altered immunological or  
 CC haematological function, in CHMR3, CHMR4 or CHMR5 for altered central and  
 CC peripheral nervous system function. The present sequence represents a  
 CC polymorphic DNA sequence of the invention

SO Sequence 1537 BP; 394 A; 326 C; 303 G; 514 T; 0 U; 0 Other;

Query Match 4.6%; Score 35.4; DB 6; Length 1537;  
 Best Local Similarity 46.8%; Pred. No. 4.7;  
 Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 522 CCGCTCAGCAAAACATGATGATCTGCTGATGAAAGCCCAACTTTCAGCCGACG 581  
 DB 913 CAGCCAGGCCCAACATGATGAAAGCCCATCTTACTTAAACCCATCTTACTTAAAAAAA 854  
 QY 582 AAGTATCTGCGAGAGAGAAAGGCTGTGTGTGCGACGCTGTCAGATTCATCTGCT 641  
 DB 853 TAAAAAATTAGCAACAGCAGAGGCGGTGCTGCAATCCAGCTACTGTAAGACTGAG 794  
 QY 642 GCGACAACCTGTGTGACACGCGAGATGCTGCGAGTGTTCGCCACGAGTGTCCGCTGCC 701  
 DB 793 GCAGGAGATGCTTAAATCCGCGAGGCGAGGTGCAATGAGCAAGATTCGCGACATG 734  
 QY 702 CATGAAGCGAAAGAAATTTGTGATCACCCGAAATTAATATTCATGCAAAAAAAA 758  
 DB 733 CACTGTAGCTGGGCGACAGAGAGACTTATCTCAAAAAAAGAAAAA 677

RESULT 7  
 ABS96867/c  
 ID ABS96867 standard; DNA; 7769 BP.  
 AC ABS96867;  
 XX  
 XX 13-AUG-2002 (first entry)  
 XX

DE Gene #3365 used to diagnose liver cancer.  
 XX Gene, liver cancer; ds, hepatocellular carcinoma; hepatotropic;  
 KM metastatic liver tumor; cytostatic; expression profile; disease state;  
 KM disease progression; drug toxicity; drug efficacy; drug metabolism.  
 XX Homo sapiens.  
 OS  
 XX WO200229103-A2.  
 PN  
 XX 11-APR-2002.  
 PD  
 XX 02-OCT-2001; 2001WO-US030589.  
 PF  
 XX 02-OCT-2000; 2000US-0237054P.  
 PR  
 XX (GENE-) GENE LOGIC INC.  
 PA  
 XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JS;  
 PI WPI; 2002-426119/45.  
 DR  
 XX  
 XX Diagnosing and detecting the progression of liver cancer, hepatocellular  
 PT carcinoma or metastatic liver tumor in a patient, involves detecting the  
 PT level of expression of two or more genes in a liver tissue sample.  
 PS Claim 1; SEQ ID NO 3365; 299pp; English.  
 XX  
 XX The invention relates to a novel method for diagnosing and detecting the  
 CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
 CC tumor in a patient, and differentiating metastatic liver cancer from  
 CC hepatocellular carcinoma in a patient, involving detecting the level of  
 CC expression of two or more genes represented in ABN93503-ABN97455 in a  
 CC tissue sample. The method of the invention has hepatotropic, and  
 CC cytostatic activity. The method is useful for diagnosing and detecting  
 CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
 CC liver carcinoma in a patient. The method is useful for identifying  
 CC expression profiles which serve as useful diagnostic markers as well as  
 CC markers that can be used to monitor disease states, disease progression,  
 CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 CC Sequence 7769 BP; 2138 A; 1595 C; 1726 G; 2306 T; 0 U; 4 Other;  
 SQ  
 Query Match 4.6%; Score 35.4; DB 6; Length 7769;  
 Best Local Similarity 46.8%; Pred. No. 10;  
 Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 522 CCGCTGACGAAACATGTGATCTCTGCGTGTATGGAAGCCCAACTTGCACCGAAGC 581  
 DB 7145 CCAGCCAGCCCAACATGTGATCTCTCTACTTAAACCCCATCTCTACTTAAAAAAA 7086  
 QY 582 AAGTTATCTGCGAGAAAGAGCTTGTGTGCGACCGCTCGTCAGATTCTTCTGCT 641  
 DB 7085 TAAAAAATTAGCCAGACGAGTGGCTGTGCTGCAATCCAGCTACTGCTGAGACTGAG 7026  
 QY 642 GCGACAACTGTGTGACACGCGAGATGCTGCGAGTGTGCGCCACGAGTGTGGCGCTGC 701  
 DB 7025 GCGAGGAATTTGCTTAAATCCGGGAGCGGAGGTTCAATGAGCAAGTTGGCGCACTG 6966  
 QY 702 CATGAAGCGAAAGAAATTGTGTGTACCCCGAAATTAATTTATGCAACAAAAA 758  
 DB 6965 CACTGTAGCTGGCGCAGAGAGAGACTTATCTCAAAAAAAGAAAAA 6909

RESULT 8  
 ABD32627\_0/c LOCUS ABD32627 Accession Abd32627  
 WP Sequence split into 5 fragments  
 WP Fragment Name Begin End  
 WP ABD32627\_0 1 110000  
 WP ABD32627\_1 100001 210000

WP ABD32627\_2 200001 310000  
 WP ABD32627\_3 300001 410000  
 WP ABD32627\_4 400001 421609  
 ID ABD32627 standard; DNA; 421609 BP.  
 AC  
 XX ABD32627;  
 XX  
 XX 18-NOV-2004 (first entry)  
 DT  
 XX Mouse cancer-associated genomic DNA MDJ3-017.  
 DE  
 XX Mouse, ds; cancer-associated protein; gene; cytostatic; cancer;  
 KM leukaemia; lymphoma; CAP.  
 KM  
 XX Mus musculus.  
 OS  
 XX WO2004074320-A2.  
 PN  
 XX 02-SEP-2004.  
 PD  
 XX 17-FEB-2004; 2004WO-US004730.  
 PF  
 XX  
 XX 14-FEB-2003; 2003US-00367094.  
 PR 14-MAR-2003; 2003US-0038838.  
 PR 15-APR-2003; 2003US-00417375.  
 PR 13-JUN-2003; 2003US-00461862.  
 PR 15-SEP-2003; 2003US-00663431.  
 PR 15-DEC-2003; 2003US-00737318.  
 PA (SAGRE-) SAGRES DISCOVERY INC.  
 PI Morris DW, Morris DW, Malandro MS;  
 PI WPI; 2004-652914/63.  
 DR  
 XX  
 XX New isolated cancer-associated polynucleotides and polypeptides useful  
 PT for diagnosing, preventing or treating cancers, especially lymphoma and  
 PT leukemia, or in screening for agents that modulate cancer.  
 PS disclosure; seqid 148; 310pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid comprising at least 10  
 CC contiguous nucleotides of any of the 233 polynucleotide sequences given  
 CC in the specification, or its complement. The nucleic acids encode cancer-  
 CC associated proteins. Also included are an expression vector comprising  
 CC the isolated nucleic acid cited above, a host cell comprising the above  
 CC recombinant nucleic acid or expression vector, a microarray for detecting  
 CC a cancer-associated (CA) nucleic acid comprising at least one probe  
 CC comprising at least 10 contiguous nucleotides of any of the above-  
 CC mentioned nucleotide sequences, an isolated polypeptide (encoded within  
 CC an open reading frame of a CA sequence selected from any of the 95  
 CC polynucleotide sequences as mentioned in the specification, or its  
 CC complement), an isolated antibody, (or its antigen binding fragment) that  
 CC binds to the above polypeptide, a hydridoma that produces the above  
 CC monoclonal antibody, a pharmaceutical composition comprising the above  
 CC antibody and a pharmaceutical excipient, a kit for detecting cancer  
 CC cells (comprising the antibody cited above, methods for diagnosing cancer  
 CC or for detecting the presence or absence of cancer cells in an  
 CC individual), a method for inhibiting growth of cancer cells in an  
 CC individual, a method for delivering a therapeutic agent to cancer cells  
 CC in an individual, an electronic library comprising the above  
 CC polynucleotide or polypeptide (or their fragments), methods of screening  
 CC for anticancer activity or for a bioactive agent capable of modulating  
 CC the activity of a CA protein (CAP), methods for detecting cancer  
 CC associated with expression of a polypeptide in a test cell sample, a  
 CC method for treating cancers and a method for inhibiting the expression of  
 CC CA gene in a cell. The composition and methods are useful for detecting,  
 CC diagnosing, preventing and treating cancers, especially lymphoma and  
 CC leukaemia. These may also be used in screening for agents that modulate  
 CC cancer. The present sequence is a mouse CAP genomic sequence. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX	Sequence	421609 BP;	107858A;	90251C;	92730G;	119218T;	0U;	115520Cher;
Qy	Query Match	4.5%;	Score 34.6;	DB 13;	Length 110000;			
	Best Local Similarity	50.3%;	Pred. No. 69;					
	Matches 85;	Conservative	0;	Mismatches	84;	Indels	0;	Gaps 0;
Qy	509	CATAGATCTGCCCCGCTCAGCAAAACATGTAATCTCGTGTATGGAAGCCCCAAA	568					
Db	65832	CCTAAGCATGGGGCCCGTGGCCATGGGTACCCATGACCGGTAAAGCTGTGCATTT	65773					
Qy	569	CTTGCACGGAACGAAGTTATCTGGCAGAGAAGAAAGCTGTGTGTCGACGCTCGATCA	628					
Db	65772	AGAGGAACGAACATAAACAGGAGCCACGGGGAGATGGGCATTAATTGCACAGGCTTATCCA	65713					
Qy	629	GATTCATTCTGCTGCGCAACCTGTGTGACACCGCGAATGCTGCGAGTG	677					
Db	65712	GCCCTCATTTTATTAAACCAACGGGGGAATGTGGAAACCCACATGTG	65664					

RESULT 9  
 ACN44398/c  
 ID ACN44398 standard; DNA; 194883 BP.  
 XX  
 AC ACN44398;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Human genomic sequence hCG20056.  
 XX  
 KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003073826-A2.  
 XX  
 PD 12-SEP-2003.  
 XX  
 PF 28-FEB-2003; 2003WO-US006235.  
 XX  
 PR 01-MAR-2002; 2002US-00087192.  
 XX  
 PA (SAGR-) SAGRES DISCOVERY.  
 XX  
 PI Morris DW;  
 XX  
 DR WPI; 2003-328604/31.  
 XX  
 PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
 PT comprises a nucleotide sequence.  
 XX  
 PS Claim 1; SEQ ID NO 826; opp; English.  
 CC The present invention relates to novel DNA and protein sequences which  
 CC are associated with carcinomas. The sequences are useful for: (i) for  
 CC screening drug candidates; (ii) for screening of bioactive agent capable  
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for  
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the  
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of  
 CC carcinoma including lymphoma. The present sequence is one such CA coding  
 CC sequence. Note: This patent is an equivalent to basic patent  
 CC US0002182586A1, for which no sequence data was published  
 CC  
 XX  
 Sequence 194883 BP; 41041 A; 50809 C; 53834 G; 49199 T; 0 U; 0 Other; .  
 Query Match 4.5%; Score 34.6; DB 11; Length 194883;  
 Best Local Similarity 50.9%; Pred. No. 92;  
 Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0

QY	150	ATATAGACCTTGGAAATGGTTTATATAA	CTGAGCACTAATATCAAGCGGTCATTCTTG	209
Db	79227	AGATGGACAGCTGCTGTAGTCCACGACT	CTCAGAGAGAGAGATGAGAGATCAGTG	79168
QY	210	GTCCAGCTCAAGAACATGTACAAAGTGA	CTGGAGTATTCGAATTGGAGAGTAGACAGAC	269
Db	79167	GAGCCACAGAGATTCAAGAGCTGAAGTGC	GCTATATGTCATCTGTGAAATATGACCATGTAC	79108
QY	270	AACGATTGGCGCATGTAATGATCCCTCA	TCCGTAATATAC	310
Db	79107	TCGAGCCTGG3CAACACAGCGAAACCT	CTATTAGAAAATTC	79067
RESULT 10				
ID	ABQ24108/c			
AC	ABQ24108 standard; DNA; 1003 BP.			
AC	ABQ24108;			
DT	12-JUL-2002 (first entry)			
DE	Oligonucleotide for detecting cytosine methylation SEQ ID NO 10699.			
KW	Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;			
KW	drug; side effect; cancer; central nervous system; cardiovascular;			
KW	gastrointestinal; respiratory system; single nucleotide polymorphism;			
KW	SNP; cell differentiation; db.			
OS	Homo sapiens.			
PN	MO200218632-A2.			
PD	07-MAR-2002.			
PP	01-SEP-2001; 2001MO-EP010074.			
PR	01-SEP-2000; 2000DE-01043826.			
PR	05-SEP-2000; 2000DE-01044543.			
XX	(EPIC-) EPIDENOMICS AG.			
XX	Olek A, Piepenbrock C, Berlin K, Guetig D;			
XX	WPI; 2002-371829/40.			
XX	Determining the degree of cytosine methylation in genomic DNA, useful for			
XX	diagnosis and prognosis, comprises selective hybridization of amplicons			
XX	from chemically treated DNA.			
XX	Claim 12; 56pp + Sequence Listing; 56pp; German.			
XX	This invention describes a novel method for determining the degree of			
XX	methylation of a particular cytosine in a motif 5'-CpG-3', present in a			
XX	genomic sample of DNA. The sample is treated chemically to convert			
XX	cytosine (C) but not methylated C, to uracil, then part of the genomic			
XX	DNA that contains the target C is amplified to form a labeled amplicon.			
XX	The amplicon is hybridised to two classes, each with at least one member,			
XX	of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the			
XX	degree of hybridisation to both classes is determined from the label on			
XX	the amplicon. From the ratio of labels hybridised to the two classes of			
XX	oligomers, the degree of methylation is calculated. The method is used:			
XX	(i) for diagnosis and/or prognosis of side effects of therapeutic drugs			
XX	and of a wide range of diseases, e.g. cancer, disorders of the central			
XX	nervous, cardiovascular, gastrointestinal and respiratory systems etc.,			
XX	particularly by detecting mutations or single nucleotide polymorphisms			
XX	(SNP's); and (ii) for differentiation of cell or tissue types and for			
XX	investigating cell differentiation. The method allows the methylation			
XX	status of many C residues to be determined simultaneously. ABO13410-			
XX	ABO54121 represent genomic DNA sequences used to illustrate the method			
XX	for determining the degree of cytosine methylation described in the			
XX	disclosure of the invention			

SQ Sequence 1003 BP; 89 A; 129 C; 313 G; 472 T; 0 U; 0 Other;

Query Match 4.5%; Score 34.4; DB 6; Length 1003;

Best Local Similarity 52.0%; Pred. No. 7.9;

Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

OY 616 CGACGCTGTCGACATTCATTCGTCGACACCTGTGTGACACGACGATGCTGCGAG 675  
 DB 986 CGCGTCCCGTCGCAACCCCTACTACGACGACTATACCAATTAATTAATCCCAA 927  
 OY 676 TGTTCGCCACCAAGTGTGCGCGTCGACATGAAGCAAAATTTGTATGTCACCCGCA 735  
 DB 926 CGCTGCCCGTCGCAACCTTAAACGACGAAAAACCAATTAATAAGAAAAAATCGAA 867  
 OY 736 TAAATATTCATGCAAAAAA 763  
 DB 866 TACATTAACCCGAAAAAAGCAAAA 839

RESULT 11

ABO24109  
 ID ABO24109 standard; DNA; 1003 BP.

AC ABO24109;

DT 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 10700.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 XX drug; side effect; cancer; central nervous system; cardiovascular;  
 XX gastrointestinal; respiratory system; single nucleotide polymorphism;  
 XX SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2.

PD 07-MAR-2002.

PF 01-SEP-2001; 2001WO-EP010074.

PR 01-SEP-2000; 2000DE-01043826.

PR 05-SEP-2000; 2000DE-01044543.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guettig D;

DR WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly for detecting mutations of single nucleotide polymorphisms  
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation

CC status of many C residues to be determined simultaneously. ABO13410-  
 CC ABO54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention

SQ Sequence 1003 BP; 472 A; 313 C; 129 G; 89 T; 0 U; 0 Other;

Query Match 4.5%; Score 34.4; DB 6; Length 1003;

Best Local Similarity 52.0%; Pred. No. 7.9;

Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

OY 616 CGACGCTGTCGACATTCATTCGTCGACACCTGTGTGACACGACGATGCTGCGAG 675  
 DB 18 CGCGTCCCGTCGCAACCCCTACTACGACGACTATACCAATTAATTAATCCCAA 77  
 OY 676 TGTTCGCCACCAAGTGTGCGCGTCGACATGAAGCAAAATTTGTATGTCACCCGCA 735  
 DB 78 CGCTGCCCGTCGCAACCTTAAACGACGAAAAACCAATTAATAAGAAAAAATCGAA 137  
 OY 736 TAAATATTCATGCAAAAAA 763  
 DB 138 TACATTAACCCGAAAAAAGCAAAA 165

RESULT 12

ABD32806.1  
 Continuation (2 of 7) of ABD32806 from base 100001 (Human cancer-associated genomic DNA ;

WP Sequence Split into 7 fragments LOCUS ABD32806 Accession Abd32806

WP	Fragment Name	Begin	End
WP	ABD32806_0	1	110000
WP	ABD32806_1	100001	210000
WP	ABD32806_2	200001	310000
WP	ABD32806_3	300001	410000
WP	ABD32806_4	400001	510000
WP	ABD32806_5	500001	610000
WP	ABD32806_6	600001	653458

Query Match 4.5%; Score 34.2; DB 13; Length 110000;

Best Local Similarity 58.3%; Pred. No. 93;

Matches 60; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 283 ATGCAATGATCCCTACCGATTAATACGACCTGCTCAAAATTCGTAGATGCTGTA 342  
 DB 43817 ATGGAATATTGAGTTCTGATGAGTGACAGTGTCTGAATTAATGACGACGACGA 43876  
 OY 343 CTTCAAGACAGTGAAGAGAGACAGTCTGCAAGATATCG 385  
 DB 43877 GGCCAATTCCTAATTAAGAACCAATTCCTGCTTCAAGATAG 43919

RESULT 13

ABL10900/C  
 ID ABL10900 standard; cDNA; 3457 BP.

AC ABL10900;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 27182.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.  
PA Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI, 2001-656860/75.  
DR P-PSDB, ABB66797.  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX Claim 1; SEQ ID NO 27182; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signaling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
XX AB12072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 3457 BP; 972 A; 732 C; 768 G; 985 T; 0 U; 0 Other;  
Query Match 4.5%; Score 34; DB 4; Length 3457;  
Best Local Similarity 51.3%; Pred. No. 19;  
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 122 CTCGATTTTCACATGATGTCGTGGAATATAGCACTTGGAATGTTGATTAACCTGG 181  
DB 2931 CACGTAATCTCACTCAAGAGATGCTACTTACCTTTGAGATGATTTGCAATGCAAG 2872  
QY 182 ACAGTAATGCAAGCGCGTCATTTCTTGTCAGCTCAGAACATGTACAAAGTGAAGCTGG 241  
DB 2871 GACGTAAATCCGCGACGCGGCAATTCATTCAGGAAAGTATTAGACATTCGAAAGA 2812  
QY 242 GATTGCACTTGAAGAAGTACGACACACAGA 275  
DB 2811 TATTGAAATCTGTAAACATTCGACGCGACGACA 2778  
RESULT 14  
AB18190/c  
ID ABL18190 standard; DNA; 4092 BP.  
XX  
XX ABL18190;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 6043.  
DB  
XX  
XX Drosophila; developmental biology; cell signaling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
PF  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-00614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.

XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell  
XX PT interactions.  
XX  
XX Claim 1; SEQ ID NO 6043; 21bp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signaling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention  
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
XX sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
XX AB12072). The sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 4092 BP; 1185 A; 876 C; 895 G; 1136 T; 0 U; 0 Other;  
Query Match 4.5%; Score 34; DB 4; Length 4092;  
Best Local Similarity 51.3%; Pred. No. 21;  
Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;  
QY 122 CTCGATTTTCACATGATGTCGTGGAATATAGCACTTGGAATGTTGATTAACCTGG 181  
DB 1046 CACGTAATCTCACTCAAGAGATGCTACTTACCTTTGAGATGATTTGCAATGCAAG 987  
QY 182 ACAGTAATGCAAGCGCGTCATTTCTTGTCAGCTCAGAACATGTACAAAGTGAAGCTGG 241  
DB 986 GACGTAAATCCGCGACGCGGCAATTCATTCAGGAAAGTATTAGACATTCGAAAGA 927  
QY 242 GATTGCACTTGAAGAAGTACGACACACAGA 275  
DB 926 TATTGAAATCTGTAAACATTCGACGCGACGACA 893  
RESULT 15  
AB113492/c  
ID ABL13492 standard; CDNA; 7160 BP.  
XX  
XX ABL13492;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DT  
XX  
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 34958.  
DB  
XX  
XX Drosophila; developmental biology; cell signaling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
XX Drosophila melanogaster.  
OS  
XX  
XX WO200171042-A2.  
PN  
XX  
XX 27-SEP-2001.  
PD  
XX  
XX 23-MAR-2001; 2001WO-US009231.  
PF  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
PR  
XX  
XX 11-JUL-2000; 2000US-00614150.  
PR  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI  
XX  
XX WPI; 2001-656860/75.  
DR  
XX  
XX P-PSDB, ABB69389.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
XX PT genes from Drosophila and for elucidating cell signaling and cell-cell  
XX PT interactions.  
XX  
XX Claim 1; SEQ ID NO 34958; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 7160 BP; 2075 A; 1468 C; 1473 G; 2144 T; 0 U; 0 Other;

Query Match 4.5%; Score 34; DB 4; Length 7160;

Best Local Similarity 51.3%; Pred. No. 28; Mismatches 75; Indels 0; Gaps 0;

Matches 79; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY	122	CTCGATTTTCA	CAATGATGTTG	CGAATATATAG	CACTTGGAAATGCTTTG	TATTAAC	TGG	181
DB	1046	CACGTATACCT	CACTCAAGAGAT	GTGTA	CTTTTCGAGTAG	TTGCA	TATGCAAG	987
QY	182	ACAGTAAATG	CAGACCGGTC	ATTCTTGGT	CCAGCTCAGACAT	GTATACAAAGT	GACTGG	241
DB	986	GACGTAAATC	CGCCAGCGGCA	TTCCAATTCA	AGCAAGTAT	TATGAGCAAT	CGAAGA	927
QY	242	GATTGCACTT	GGAAGAGTAG	CAGACAAACAG	A	275		
DB	926	TATTGAAATC	TGTAAACATT	CGAGCGCAG	CACA	893		

Search completed: March 4, 2006, 07:12:51

Job time : 594 secs





US-08-232-463-14

Query Match 5.1%; Score 39.2; DB 2; Length 7218;  
Best Local Similarity 7.7%; Pred. No. 0.12;  
Matches 32; Conservative 197; Mismatches 185; Indels 0; Gaps 0;

102 ACAGCGCAGCAGATTTTCCCGATTTTCAATGATGTTGCGAATATAGACATTG 161  
1505 AAAACGGCATGAGCATCACTGATTAATCTATATGCAAGTAGTAAAGAGTAGAA 1446  
162 GAATGTTGTAATAACTGACAGATAATGACAGCGCTCATTTGCTCAGCTCAGA 221  
1445 GAATTTGCTACRR 1386  
222 ACATGTACAAAGTGAAGTGGATTGCAACTTGAAGAAAGTACGACACACAGATTGGC 281  
1385 RRR 1326  
282 CATGCAATGATCCCTCAGCAATTAATACAGCGCTGCAAAATATCGTATGCTGT 341  
1345 RRR 1266  
342 ACTTCAAGACAGTGAAGACAGACAGTTCGCAACAAGTATCTGTATTTGGGTGAGCG 401  
1265 RRR 1206  
402 CATGCTGGGATTTATGAAAGGACAGAACTTGCAACTTTGCTTACCACTGGGCTGAC 461  
1205 RRR 1146  
462 CTCTAGCAAAACATTCGAACTATAGAAACCGAAGTGGATGTCCTCCATAGA 515  
1145 RRR 1092

## RESULT 2

US-09-949-016-45471  
; Sequence 45471, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 45471  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-45471

Query Match 4.5%; Score 34.6; DB 3; Length 601;  
Best Local Similarity 50.9%; Pred. No. 1.2; Indels 0; Gaps 0;  
Matches 82; Conservative 0; Mismatches 79;

150 ATATAGACTTGGAAATGTTGATTAATCTGACAGTAATATGACAGCGCTCATTTCTG 209  
375 AGATGGACCTGCTGCTAGTCCAGCTACTCAGAGAGAGAGTGAAGAGATCAGTG 434  
210 GTCCAGCTCAGAACATGTACAAAGTGAAGTGGATTGCAACTTGAAGAAATGACAGC 269  
435 GAGCCACGAGGTTCAAGGCTGAAGTGGCTATGATTCATCTGTGAAATGACACTGTAC 494  
270 AACAGATTGGCCATGCAATGATCCCTACCGATTAATACC 310

Db

495 TCCAGCTGGGCAACACAGCGAAACCTTACTTAAGAAATC 535

RESULT 3  
US-09-949-016-14790/c  
; Sequence 14790, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14790  
; LENGTH: 8297  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14790

Query Match 4.5%; Score 34.6; DB 3; Length 8297;  
Best Local Similarity 47.8%; Pred. No. 3.9;  
Matches 100; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

555 ATGAAAGCCCAACTTGCACCGAAGTATCTGCGACAGAAAGAAAGCTGTGTGT 614  
7059 ATGAAAGAACCTGCTCTTACTTAATAACAAATTTAGTGGCTGTGCTGTCT 7000  
615 GCGACGCTGCTGCAGATTCATTTGCTGCGACCAACTGTGTGACACGGAGATGCTGCA 674  
6999 GTATCCAGCTACTCAGAGAGGCTGAGGACAGAAATGCTTGAATGAAAGGACAGAG 6940  
675 GTGTTGCGCACCACTGTTCCGCTGCGCATGAGCGAAAGAAATTTGATGACACCCCGA 734  
6939 TTGTGTGAACCGAGATCACCGCTTCCCTTCAGCTGGGCAACAAAGCGAACTCGT 6880  
735 ATAAATATTCATGCAAAAAA 763  
6879 CTCAAAAA 6851

## RESULT 4

US-09-949-016-12733/c  
; Sequence 12733, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12733  
; LENGTH: 17883  
; TYPE: DNA  
; ORGANISM: Human



[illegible]

```

RESULT 8
US-09-949-016-15779
/ Sequence 15779, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15779
/ LENGTH: 265038
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)-(265038)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15779

```

Query Match	4.5%;	Score 34;	DB 3;	Length 265038;
Best Local Similarity	63.4%;	Pred. No. 30;		
Matches 52;	Conservative	0;	Mismatches 30;	Indels 0;
			Gaps	0;
QY	156	CACTGGAAATGTTTGATTAAC	TGGAAGTAAATGACAGCGGTCATTTCTGTGCAG	215
DB	168003	CAC TGGAAAGATGATTGATTA	CTGTGCGCACTCACTGAGAGAGGGTTTTTTTTGGACAG	168062
QY	216	CTCAGAACATGTCAAAATGGA		237
DB	168063	CAAAAAAAAAAAAAAAAAAGGA		168084

```

RESULT 9
US-09-326-402C-1/C
; Sequence 1, Application US/09326402C
; Patent No. 6759192
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bouguenelert, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Polynorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)
; FILE REFERENCE: GEN-F112XC1
; CURRENT APPLICATION NUMBER: US/09/326,402C
; CURRENT FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 60/088,187

```

```

1 PRIOR FILING DATE: 1998-06-05
2 PRIOR APPLICATION NUMBER: 60/102,324
3 PRIOR FILING DATE: 1998-09-28
4 NUMBER OF SEQ ID NOS: 22
5 SOFTWARE: PatentIn version 3.1
6 SEQ ID NO 1
7 LENGTH: 106746
8 TYPE: DNA
9 ORGANISM: Homo sapiens
10 FEATURE:
11 NAME/KEY: misc.feature
12 LOCATION: 1..68647
13 OTHER INFORMATION: 5'regulation region
14 FEATURE:
15 NAME/KEY: misc.feature
16 LOCATION: 66647..68647
17 OTHER INFORMATION: promoter
18 FEATURE:
19 NAME/KEY: misc.feature
20 LOCATION: 97156..106746
21 OTHER INFORMATION: 3'regulation region
22 FEATURE:
23 NAME/KEY: exon
24 LOCATION: 68648..68741
25 OTHER INFORMATION: exon0
26 FEATURE:
27 NAME/KEY: exon
28 LOCATION: 70647..70794
29 OTHER INFORMATION: exon1
30 FEATURE:
31 NAME/KEY: exon
32 LOCATION: 82208..82296
33 OTHER INFORMATION: exon2
34 FEATURE:
35 NAME/KEY: exon
36 LOCATION: 83613..83823
37 OTHER INFORMATION: exon3
38 FEATURE:
39 NAME/KEY: exon
40 LOCATION: 85298..85417
41 OTHER INFORMATION: exon4
42 FEATURE:
43 NAME/KEY: exon
44 LOCATION: 86389..86445
45 OTHER INFORMATION: exon5
46 FEATURE:
47 NAME/KEY: exon
48 LOCATION: 87496..87522
49 OTHER INFORMATION: exon6
50 FEATURE:
51 NAME/KEY: exon
52 LOCATION: 87650..87775
53 OTHER INFORMATION: exon6bis
54 FEATURE:
55 NAME/KEY: exon
56 LOCATION: 88295..88383
57 OTHER INFORMATION: exon7
58 FEATURE:
59 NAME/KEY: exon
60 LOCATION: 89484..89649
61 OTHER INFORMATION: exon8
62 FEATURE:
63 NAME/KEY: exon
64 LOCATION: 92749..97155
65 OTHER INFORMATION: exon9
66 FEATURE:
67 NAME/KEY: exon
68 LOCATION: 92749..92883
69 OTHER INFORMATION: exon9bis
70 FEATURE:
71 NAME/KEY: exon
72 LOCATION: 95821..97155
73 OTHER INFORMATION: exon9ter

```

```

1 PRIOR FILING DATE: 1998-06-05
2 PRIOR APPLICATION NUMBER: 60/000000
3 PRIOR FILING DATE: 1998-05-22
4 NUMBER OF SEQ ID NOS: 22
5 SOFTWARE: PatentIn version 3.0.1
6 SEQ ID NO 1
7 LENGTH: 106746
8 TYPE: DNA
9 ORGANISM: Homo sapiens
10 FEATURE:
11 NAME/KEY: misc.feature
12 LOCATION: 1..68647
13 OTHER INFORMATION: 5'regulatory
14 FEATURE:
15 NAME/KEY: misc.feature
16 LOCATION: 66647..68647
17 OTHER INFORMATION: promoter
18 FEATURE:
19 NAME/KEY: misc.feature
20 LOCATION: 97156..106746
21 OTHER INFORMATION: 3'regulatory
22 FEATURE:
23 NAME/KEY: exon
24 LOCATION: 68648..68741
25 OTHER INFORMATION: exon0
26 FEATURE:
27 NAME/KEY: exon
28 LOCATION: 70647..70794
29 OTHER INFORMATION: exon1
30 FEATURE:
31 NAME/KEY: exon
32 LOCATION: 82208..82296
33 OTHER INFORMATION: exon2
34 FEATURE:
35 NAME/KEY: exon
36 LOCATION: 83613..83823
37 OTHER INFORMATION: exon3
38 FEATURE:
39 NAME/KEY: exon
40 LOCATION: 85298..85417
41 OTHER INFORMATION: exon4
42 FEATURE:
43 NAME/KEY: exon
44 LOCATION: 86389..86445
45 OTHER INFORMATION: exon5
46 FEATURE:
47 NAME/KEY: exon
48 LOCATION: 87496..87522
49 OTHER INFORMATION: exon6
50 FEATURE:
51 NAME/KEY: exon
52 LOCATION: 87650..87775
53 OTHER INFORMATION: exon6bis
54 FEATURE:
55 NAME/KEY: exon
56 LOCATION: 88295..88383
57 OTHER INFORMATION: exon7
58 FEATURE:
59 NAME/KEY: exon
60 LOCATION: 89484..89649
61 OTHER INFORMATION: exon8
62 FEATURE:
63 NAME/KEY: exon
64 LOCATION: 92749..97155
65 OTHER INFORMATION: exon9
66 FEATURE:
67 NAME/KEY: exon
68 LOCATION: 92749..92883
69 OTHER INFORMATION: exon9bis
70 FEATURE:
71 NAME/KEY: exon
72 LOCATION: 95821..97155
73 OTHER INFORMATION: exon9ter

```

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 70647..70794  
OTHER INFORMATION: homology with genset EST : A241850  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 68648..68741  
OTHER INFORMATION: homology with genset EST : A241850  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 82205..82229  
OTHER INFORMATION: homology with genset EST : A241850  
FEATURE:  
NAME/KEY: allele  
LOCATION: 278  
OTHER INFORMATION: 99-1601-278 : polymorphic base A or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 402  
OTHER INFORMATION: 99-1601-402 : polymorphic base A or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 472  
OTHER INFORMATION: 99-1601-472 : polymorphic base A or T  
FEATURE:  
NAME/KEY: allele  
LOCATION: 2955  
OTHER INFORMATION: 99-13801-100 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 12167  
OTHER INFORMATION: 99-13806-166 : polymorphic base G or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 12536  
OTHER INFORMATION: 99-13799-376 : polymorphic base T or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 17593  
OTHER INFORMATION: 99-13798-297 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 17606  
OTHER INFORMATION: 99-13798-284 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 22079  
OTHER INFORMATION: 99-1602-200 : polymorphic base G or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 28964  
OTHER INFORMATION: 99-13794-186 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 29003  
OTHER INFORMATION: 99-13794-147 : polymorphic base C or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 31077  
OTHER INFORMATION: 99-13812-384 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 31766  
OTHER INFORMATION: 99-13805-313 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 34791  
OTHER INFORMATION: 99-1587-281 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 45751  
OTHER INFORMATION: 99-1582-430 : polymorphic base C or T  
FEATURE:

NAME/KEY: allele  
LOCATION: 49847  
OTHER INFORMATION: 99-1585-465 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49855  
OTHER INFORMATION: 99-1585-457 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49866  
OTHER INFORMATION: 99-1585-426 : polymorphic base G or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49900  
OTHER INFORMATION: 99-1585-412 : polymorphic base G or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49906  
OTHER INFORMATION: 99-1585-406 : polymorphic base C or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49921  
OTHER INFORMATION: 99-1585-391 : polymorphic base C or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 49939  
OTHER INFORMATION: 99-1585-373 : polymorphic base G or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 50256  
OTHER INFORMATION: 99-1585-55 : polymorphic base C or A  
FEATURE:  
NAME/KEY: allele  
LOCATION: 54955  
OTHER INFORMATION: 99-1607-373 : polymorphic base T or C  
FEATURE:  
NAME/KEY: allele  
LOCATION: 64239  
OTHER INFORMATION: 99-1577-105 : polymorphic base A or G  
FEATURE:  
NAME/KEY: allele  
LOCATION: 65436  
OTHER INFORMATION: 99-1591-235 : polymorphic base A or G

Query Match 4.4%; Score 33.4; DB 3; Length 106746;  
Best Local Similarity 51.7%; Pred. No. 30;  
Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 342 ACTTCAAGACAGTGAAGAGACAGTTCTGCAACAGTATCGTGTATTGGGTGACCG 401  
DB 91433 AGTGCAGAGCCAGTGAATTTACAGCAAGTGGCAAGAAATGACATGCTTCCGGGTGACCA 91374  
QY 402 CATCGCTGGATTATGAAAGGACGAAACTTGACCAATTTGCTTAACCAAGTGGCTGAAC 461  
DB 91373 GACTACTGCAATGAAAAGAGGACATCTTTGGCTTGAAGAAAAGAAACAAACAGAAAT 91314  
QY 462 CTCTAGCAACATTCGCAACTTATGAA 488  
DB 91313 CCTTCACAAACATCTGATGCTGCCGA 91287

RESULT 10  
US-09-326-402C-12/c  
Sequence 12, Application US/09326402C  
Patent No. 6759192  
GENERAL INFORMATION:  
APPLICANT: Blumenfeld, Marta  
APPLICANT: Boungeleret, Lydie  
APPLICANT: Chumakov, Ilya  
TITLE OF INVENTION: Polymorphic Markers of Prostate Carcinoma Tumor Antigen-1 (PCTA-1)  
FILE REFERENCE: GEN-T112K01  
CURRENT APPLICATION NUMBER: US/09/326,402C  
CURRENT FILING DATE: 1999-06-04

```
PRIOR APPLICATION NUMBER: 60/088,187
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/102,324
PRIOR FILING DATE: 1998-09-28
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 106746
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..68647
OTHER INFORMATION: 5' regulation region
FEATURE:
NAME/KEY: misc feature
LOCATION: 66647..68647
OTHER INFORMATION: promoter
FEATURE:
NAME/KEY: misc feature
LOCATION: 97156..106746
OTHER INFORMATION: 3' regulation region
FEATURE:
NAME/KEY: exon
LOCATION: 70647..70794
OTHER INFORMATION: exon1
FEATURE:
NAME/KEY: exon
LOCATION: 82208..82296
OTHER INFORMATION: exon2
FEATURE:
NAME/KEY: exon
LOCATION: 83613..83823
OTHER INFORMATION: exon3
FEATURE:
NAME/KEY: exon
LOCATION: 85298..85417
OTHER INFORMATION: exon4
FEATURE:
NAME/KEY: exon
LOCATION: 86389..86445
OTHER INFORMATION: exon5
FEATURE:
NAME/KEY: exon
LOCATION: 87496..87522
OTHER INFORMATION: exon6
FEATURE:
NAME/KEY: exon
LOCATION: 87650..87775
OTHER INFORMATION: exon6bis
FEATURE:
NAME/KEY: exon
LOCATION: 88295..88383
OTHER INFORMATION: exon7
FEATURE:
NAME/KEY: exon
LOCATION: 89484..89649
OTHER INFORMATION: exon8
FEATURE:
NAME/KEY: exon
LOCATION: 92749..97155
OTHER INFORMATION: exon9
FEATURE:
NAME/KEY: exon
LOCATION: 92749..92883
OTHER INFORMATION: exon9bis
FEATURE:
NAME/KEY: exon
LOCATION: 95821..97155

OTHER INFORMATION: exon9ter
FEATURE:
NAME/KEY: misc feature
LOCATION: 70647..70794
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: misc feature
LOCATION: 68648..68741
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: misc feature
LOCATION: 82208..82229
OTHER INFORMATION: homology with genset EST : A241850
FEATURE:
NAME/KEY: allele
LOCATION: 278
OTHER INFORMATION: 99-1601-278 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 402
OTHER INFORMATION: 99-1601-402 : polymorphic base w= A or T
FEATURE:
NAME/KEY: allele
LOCATION: 472
OTHER INFORMATION: 99-1601-472 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 2955
OTHER INFORMATION: 99-13801-100 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 12167
OTHER INFORMATION: 99-13806-166 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 12536
OTHER INFORMATION: 99-13799-376 : polymorphic base T or G
FEATURE:
NAME/KEY: allele
LOCATION: 17593
OTHER INFORMATION: 99-13798-297 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 17606
OTHER INFORMATION: 99-13798-284 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 22079
OTHER INFORMATION: 99-1602-200 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 28964
OTHER INFORMATION: 99-13794-186 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 29003
OTHER INFORMATION: 99-13794-147 : polymorphic base C or G
FEATURE:
NAME/KEY: allele
LOCATION: 31077
OTHER INFORMATION: 99-13812-384 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 31766
OTHER INFORMATION: 99-13805-313 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 34791
OTHER INFORMATION: 99-1587-281 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 45751
OTHER INFORMATION: 99-1582-430 : polymorphic base C or T
```

```
FEATURE:
NAME/KEY: allele
LOCATION: 49847
OTHER INFORMATION: 99-1585-465 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 49855
OTHER INFORMATION: 99-1585-457 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 49886
OTHER INFORMATION: 99-1585-426 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 49900
OTHER INFORMATION: 99-1585-412 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 49921
OTHER INFORMATION: 99-1585-391 : polymorphic base C or A
FEATURE:
NAME/KEY: allele
LOCATION: 49939
OTHER INFORMATION: 99-1585-373 : polymorphic base G or A
FEATURE:
NAME/KEY: allele
LOCATION: 50256
OTHER INFORMATION: 99-1585-55 : polymorphic base C or A
FEATURE:
NAME/KEY: allele
LOCATION: 54955
OTHER INFORMATION: 99-1607-373 : polymorphic base T or C
FEATURE:
NAME/KEY: allele
LOCATION: 64239
OTHER INFORMATION: 99-1577-105 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 65436
OTHER INFORMATION: 99-1591-235 : polymorphic base A or G
```

```
Query Match 4.4%; Score 33.4; DB 3; Length 106746;
Best Local Similarity 51.7%; Pred. No. 30;
Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
```

```
Qy 342 ACTTCAAGACAGTGAAGAGACAGTCTGCAACAGTATCGTATTGGGTGAGCG 401
Db 91433 ACTGCAAGCCAGTGAATTACAGCAAGTGCAGAAAGAAATGACATGCTCCGGTGAGCA 91374
Qy 402 CATCGCTGGATTATGAAGGACGAACTTGACCAATTGCTAACCAAGTGGCTGAAC 461
Db 91373 GACTACTGGAATGAAAAGAGCAGCATCTTGTGCTAGAAAACAGAAACAGAAAT 91314
Qy 462 CTCTAGCAAACTTGCACAACTTAAAG 488
Db 91313 CCTTCACAAACATCTGAGCTGCCGAA 91287
```

```
RESULT 11
US-09-620-312D-351
Sequence 351, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Aundl, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
```

```
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Jang
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 351
LENGTH: 5238
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (278) .. (1108)
```

```
Query Match 4.3%; Score 32.6; DB 3; Length 5238;
Best Local Similarity 57.3%; Pred. No. 14;
Matches 59; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
```

```
Qy 343 CTTCAAGACAGTGAAGAGACAGTCTGCAACAGTATCGTATTGGGTGAGCGC 402
Db 38 CATCAAGATGAAGAGAGAGACAGTCTTTCGAGAGTGATGCTTTCCCGGAGACC 97
Qy 403 ATCGCTGGATTATGAAGGACGAACTTGACCAATTGCT 445
Db 98 CTTGCCAGATATATGACAGAGACACACCAACCAATCAGCT 140
```

```
RESULT 12
US-08-678-304-2
Sequence 2, Application US/08678304
Patent No. 5663067
GENERAL INFORMATION:
APPLICANT: Xu, Shuang-yong
APPLICANT: Xiao, Jian-ping
APPLICANT: Maunus, Robert B.
TITLE OF INVENTION: METHOD FOR CLONING AND
TITLE OF INVENTION: PRODUCING THE SAPI RESTRICTION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: NEW ENGLAND BIOLABS, INC.
STREET: 32 TOZER ROAD
CITY: BEVERLY
STATE: MA
COUNTRY: US
ZIP: 01915
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/678,304
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
```

```
NAME: WILLIAMS, GREGORY D
REGISTRATION NUMBER: 30901
REFERENCE/DOCKET NUMBER: NBB-123
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-927-5054
TELEFAX: 508-927-1705
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1302 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1299
US-08-678-304-2

Query Match      4.2%; Score 32.2; DB 2; Length 1302;
Best Local Similarity 54.7%; Pred. No. 9.7;
Matches 64; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 163 AATGTTTGAATACTGGACAGTAATGACAGCGGTCATTCTTGTCAGCTCAGAA 222
DB 645 AATGATATCATCAACCGAAGGTTGATGTGGCGGATTCTTCTGATAGGTTCAAGC 704
QY 223 CATGTACAAGTGAGCTGGGATTGCACTTGGAAGAGTAGACAGCAACAGATTGC 279
DB 705 AATGCTGACGATGTCGACGAGTCCCTTGGAGCGGATGATCATATTATGTTTC 761

RESULT 13
US-09-641-638-232
Sequence 232, Application US/09641638
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bouguerelet, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GENSET.051CPI
CURRENT APPLICATION NUMBER: US/09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 232
LENGTH: 420
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 198
OTHER INFORMATION: 10-94-198 : polymorphic base G or T
NAME/KEY: misc_binding
LOCATION: 178..197
OTHER INFORMATION: 10-94-198.misl, potential
NAME/KEY: misc_binding
LOCATION: 199..218
OTHER INFORMATION: 10-94-198.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 1..19
OTHER INFORMATION: upstream amplification primer
```

```
NAME/KEY: primer_bind
LOCATION: 403..420
OTHER INFORMATION: downstream amplification primer, complement
NAME/KEY: misc_binding
LOCATION: 186..210
OTHER INFORMATION: 10-94-198 potential probe
NAME/KEY: misc_feature
LOCATION: 351
OTHER INFORMATION: n=a, g, c or t
US-09-641-638-232

Query Match      4.2%; Score 32; DB 3; Length 420;
Best Local Similarity 54.4%; Pred. No. 6.7;
Matches 62; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

QY 87 GCTTAGCGAAGTACGAGCGCGAGATTTCTCGATTTTCACAAATGATGTCGTC 146
DB 173 GTCCATTCGAGGTTGAATAGCTAGKCTTTTCTTTTTCATTAATAAGCACTT 232
QY 147 GAATATATGACATTTGAATGTTGATTAACCTGACAGTAATGACAGCGCG 200
DB 233 AACTTAAAGTGCTTAATGCTTTGTAAGCTGAGATCTAATGAGGACAAGG 286

RESULT 14
US-10-170-097-232
Sequence 232, Application US/10170097
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bouguerelet, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
FILE REFERENCE: GEN-114KC2D1
CURRENT APPLICATION NUMBER: US/10/170,097
PRIOR FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm
SEQ ID NO 232
LENGTH: 420
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 198
OTHER INFORMATION: 10-94-198 : polymorphic base G or T
NAME/KEY: misc_binding
LOCATION: 178..197
OTHER INFORMATION: 10-94-198.misl, potential
NAME/KEY: misc_binding
LOCATION: 199..218
OTHER INFORMATION: 10-94-198.mis2, potential complement
NAME/KEY: primer_bind
LOCATION: 1..19
OTHER INFORMATION: upstream amplification primer
NAME/KEY: primer_bind
LOCATION: 403..420
```



/ OTHER INFORMATION: downstream amplification primer, complement  
/ FEATURE:  
/ NAME/KEY: misc\_binding  
/ LOCATION: 186..210  
/ OTHER INFORMATION: 10-94-198 potential probe  
/ FEATURE:  
/ NAME/KEY: misc\_feature  
/ LOCATION: 351  
/ OTHER INFORMATION: n=a, g, c or t  
US-10-170-097-232

Query Match 4.2%; Score 32; DB 3; Length 420;  
Best Local Similarity 54.4%; Pred. No. 6.7;  
Matches 62; Conservative 1; Mismatches 51; Indels 0; Gaps 0;

QY 87 GTCTAGCCAAAGTGACAGCGGAGGCAATTTCTCGATTTCACATGATGTC 146  
DB 173 GTCCATTCCAAAGTTGTAATGCTAGKCTTTTTCATTAATAAGACCAT 232  
QY 147 GAAATATACACTTGAAATGTTGATTAATGACAGTAATGACAGCGG 200  
DB 233 AACTTAAGTGGTGAATGCTTTGTAAGCTGAATCTAATGGGACAAG 286

## RESULT 15

US-09-621-976-2446  
/ Sequence 2446, Application US/09621976  
/ Patent No. 6639063  
/ GENERAL INFORMATION:  
/ APPLICANT: Dumas Milne Edwards, J.B.  
/ APPLICANT: Jobert, S.  
/ APPLICANT: Giordano, J.Y.  
/ FILE REFERENCE: GENSET 054PR2  
/ CURRENT APPLICATION NUMBER: US/09/621,976  
/ NUMBER OF SEQ ID NOS: 19335  
/ SOFTWARE: Patent.pm  
/ SEQ ID NO 2446  
/ LENGTH: 430  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: CDS  
/ LOCATION: 63..254  
US-09-621-976-2446

Query Match 4.2%; Score 32; DB 3; Length 430;  
Best Local Similarity 60.2%; Pred. No. 6.7;  
Matches 53; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 GAGAACTGTAATGTCGGCGGCTGTGATGCTGTCCTGCGCCCTGTTCTCTATGC 60  
DB 7 GTGAACCTCTGAGACTGCCAGTGTGTAAGGAGACACACGCGCTTTTACACATGG 66  
QY 61 CGAAGCAGGCTTTGTTGTCGGAATAGT 88  
DB 67 CTAAACGCGATTTCTTTGACGCAAT 94

Search completed: March 4, 2006, 07:29:57  
Job time : 190 secs

***This Page Blank (uspto)***

OM nucleic - nucleic search, using SW model

Title: US-10-527-771-9

Sequence: 1 gaggaactgctatgtcggcgg.....tcatgcaaaaaaaaaa 763

Scoring table: IDENTITY\_NUC

Searched: 7218535 seqs, 1096242582 residues

Total number of hits satisfying chosen parameters: 14437070

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA New:\*

```

1: /cgn2_6/ptodatac/1/pubpna/US08_NEW_PUB.seq.*
2: /cgn2_6/ptodatac/1/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodatac/1/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodatac/1/pubpna/PCr_NEW_PUB.seq.*
5: /cgn2_6/ptodatac/1/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodatac/1/pubpna/US09_NEW_PUB.seq.*
7: /cgn2_6/ptodatac/1/pubpna/US10_NEW_PUB.seq.*
8: /cgn2_6/ptodatac/1/pubpna/US10_NEW_PUB.seq.*
9: /cgn2_6/ptodatac/1/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodatac/1/pubpna/US11_NEW_PUB.seq.*
11: /cgn2_6/ptodatac/1/pubpna/US11_NEW_PUB.seq.*
12: /cgn2_6/ptodatac/1/pubpna/US11_NEW_PUB.seq.*
13: /cgn2_6/ptodatac/1/pubpna/US60_NEW_PUB.seq.*

```

**Pred. No.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	763	100.0	763	8	US-10-527-771-9	Sequence 9, Appli
2	273.4	35.8	828	8	US-10-527-771.1	Sequence 1, Appli
3	36.6	4.6	142303	12	US-11-121-086-42	Sequence 4, Appl
4	34.8	4.6	551	6	US-09-925-0654-501602	Sequence 501602,
5	33.4	4.4	539	6	US-09-925-0654-611636	Sequence 611636,
6	33.4	4.4	538	6	US-09-925-0654-524748	Sequence 524748,
7	33.4	4.4	570	6	US-09-925-0654-652826	Sequence 652826,
8	33.4	4.4	570	6	US-09-925-0654-652827	Sequence 652827,
9	33.4	4.4	149382	8	US-10-925-561-13272	Sequence 13272, A
10	33.2	4.4	574	6	US-09-925-0654-301865	Sequence 301865,
11	33.2	4.3	570	6	US-09-925-0654-625825	Sequence 625825,
12	32.6	4.3	576	6	US-09-925-0654-193665	Sequence 193665,
13	32.6	4.3	612	6	US-09-925-0654-745205	Sequence 745205,
14	32.4	4.2	559	6	US-09-925-0654-189082	Sequence 189082,
15	32.4	4.2	591	6	US-09-925-0654-250352	Sequence 250352,
16	32.4	4.2	646	6	US-09-925-0654-700306	Sequence 700306,
17	32.4	4.2	646	6	US-09-925-0654-700307	Sequence 700307,
18	32.4	4.2	140802	8	US-10-925-561-13445	Sequence 13445, A
19	32.4	4.2	305312	8	US-10-925-561-13336	Sequence 13336, A
20	32.4	4.2	592	6	US-09-925-0654-742765	Sequence 742765,

21	32.2	4.2	634	6	US-09-925-065A-955305	Sequence 955305,
22	32.2	4.2	3158	6	US-09-925-065A-976495	Sequence 676495,
23	32	4.2	589	6	US-09-925-065A-407938	Sequence 407938,
24	32	4.2	589	6	US-09-925-065A-407939	Sequence 407939,
25	32	4.2	589	6	US-09-925-065A-407940	Sequence 407940,
26	32	4.2	79528	8	US-10-276-233A-6	Sequence 6, Appl
27	31.8	4.2	617	6	US-09-925-065A-749100	Sequence 749100,
28	31.8	4.2	2182	8	US-10-510-386-207	Sequence 207, App
29	31.8	4.2	212716	12	US-11-121-086-95	Sequence 95, Appl
30	31.6	4.1	453	6	US-09-925-065A-347042	Sequence 347042,
31	31.6	4.1	1304	12	US-11-112-908-457	Sequence 457, Appl
32	31.6	4.1	155515	12	US-11-112-908-457	Sequence 42, App
33	31.6	4.1	159660	12	US-11-112-908-43	Sequence 41, Appl
34	31.6	4.1	17762	12	US-11-112-908-41	Sequence 43, Appl
35	31.4	4.1	552	6	US-09-925-065A-46672	Sequence 46672, A
36	31.4	4.1	2069	8	US-10-689-742-139	Sequence 139, App
37	31.4	4.1	2879	9	US-11-072-512-1564	Sequence 1564, App
38	31.2	4.1	569	6	US-09-925-065A-248083	Sequence 248083,
39	31.2	4.1	670	6	US-09-925-065A-503887	Sequence 503887,
40	31	4.1	552	6	US-09-925-065A-46673	Sequence 46673, A
41	31	4.1	657	6	US-09-925-065A-34824	Sequence 34824, A
42	30.8	4.0	201	12	US-11-124-368A-1459	Sequence 1459, App
43	30.8	4.0	201	12	US-11-124-368A-1479	Sequence 1479, App
44	30.8	4.0	201	12	US-11-124-368A-13286	Sequence 13286, A
45	30.8	4.0	649	6	US-09-925-065A-143049	Sequence 143049,

```

US-10-527-771-9
RESULT 1
: Sequence 9, Application US/10527771
: Publication NO. US20050271683A1
: GENERAL INFORMATION:
: APPLICANT: University Gent
: TITLE OF INVENTION: Osetecgia vaccine
: FILE REFERENCE: 2002-015
: CURRENT APPLICATION NUMBER: US/10/527,771
: CURRENT FILING DATE: 2005-03-11
: PRIOR APPLICATION NUMBER: US 10/243,319
: PRIOR FILING DATE: 2002-09-13
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 9

```

```

; LENGTH: 763
; TYPE: DNA
; ORGANISM: Oostertagia ostertagi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(706)
;
US-10-527-771-9

```

Query Match	100.0%;	Score 763;	DB 8;	Length 763;
Best Local Similarity	100.0%;	Pred. No. 2.4e-216;		
Matches 763; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	61	CGAAGCAGCGCTTTGTTGTCGGAATATGCTTAAAGCCAAAGTAGACACGCGGAGCGAGATTTT	120
Db	61	CGAAGCAGCGCTTTGTTGTCGGAATATGCTTAAAGCCAAAGTAGACACGCGGAGCGAGATTTT	120
Qy	1	GAAATCTGGCTAATGTCGGACGGCTGTTGTAAGTGGCTATTCCTCGAGCCGTTGCTCTTAAGC	60
Db	1	GAAATCTGGCTAATGTCGGACGGCTGTTGTAAGTGGCTATTCCTCGAGCCGTTGCTCTTAAGC	60
Qy	121	CCTCGAATTTTCCAAATGATGTTGCTCGGAATATATAGCACTTGGAAATGTTGATTAACCTG	180
Db	121	CCTCGAATTTTCCAAATGATGTTGCTCGGAATATATAGCACTTGGAAATGTTGATTAACCTG	180
Qy	181	GACAGTAAATGACAGACGGCGCTCAATCTTGGTCCAGCTCAGAACCAATGTACAAAGTGGACTG	240
Db	181	GACAGTAAATGACAGACGGCGCTCAATCTTGGTCCAGCTCAGAACCAATGTACAAAGTGGACTG	240

QY	241	GGATTGCACTTTGGAAAGAAAGTACAGACCAACAAGATTGGCCATGCAATGATCCCTTACC	300
Db	241	GGATTGCAACTTTGGAAAGAAAGTACAGACCAACAAGATTGGCCATGCAATGATCCCTTACC	300
QY	301	GATAAATATCCAGCCTGGGCTCAAAATATATGCTAGATGGCTGTACTTCAAGAAGATGTAAGA	360
Db	301	GATAAATATCCAGCCTGGGCTCAAAATATATGCTAGATGGCTGTACTTCAAGAAGATGTAAGA	360
QY	361	AGAGACAGTTCTGCACAACAAGTATCGTGGTATTGGGTAGCGCATCGCTGGGATTATGAA	420
Db	361	AGAGACAGTTCTGCACAACAAGTATCGTGGTATTGGGTAGCGCATCGCTGGGATTATGAA	420
QY	421	AGGACGGAACCTTGACCAATTTGGCTTACCAAGTGGGCTGAAACCTTATGACAAACATTGCAAA	480
Db	421	AGGACGGAACCTTGACCAATTTGGCTTACCAAGTGGGCTGAAACCTTATGACAAACATTGCAAA	480
QY	481	CTATAGAAAACGGAAGGTGGATGTGATGGTCCCATAGATCTGCCCCGCTCAGCAAAAACATGTGT	540
Db	481	CTATAGAAAACGGAAGGTGGATGTGATGGTCCCATAGATCTGCCCCGCTCAGCAAAAACATGTGT	540
QY	541	AGTATCTGCTGCTGTATGTGAAAGGCCCAAACTTGCAACGAAAGAAATTATCTGGCAGAAAG	600
Db	541	AGTATCTGCTGCTGTATGTGAAAGGCCCAAACTTGCAACGAAAGAAATTATCTGGCAGAAAG	600
QY	601	AAAGCTGTGTGTGTGTCGACGCTCTGTCAGATTCAATCTGTGTCGACAACTGTGTGTACAC	660
Db	601	AAAGCTGTGTGTGTGTCGACGCTCTGTCAGATTCAATCTGTGTCGACAACTGTGTGTACAC	660
QY	661	GGGAAGTGTGTGGCAGTGTTCGCCACCAAGTGTGGGGGTGGCCATGAAAGCGAAAGAATTT	720
Db	661	GGGAAGTGTGTGGCAGTGTTCGCCACCAAGTGTGGGGGTGGCCATGAAAGCGAAAGAATTT	720
QY	721	GGTAGTCACCCCGAATTAATAATTTCATGCAAAAAAATTTTTTTTTTTT	763
Db	721	GGTAGTCACCCCGAATTAATAATTTCATGCAAAAAAATTTTTTTTTTTT	763

```

RESULT 2
US-10-527-771-1
; Sequence 1, Application US/10527771
; Publication NO. US20050271683A1
; GENERAL INFORMATION:
; APPLICANT: Universlty Gent
; TITLE OF INVENTION: Osterlagia vaccine
; FILE REFERENCE: 2002-015
; CURRENT APPLICATION NUMBER: US/10/527, 771
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US 10/243, 319
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ. ID NO 1
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Osterlagia osterlagi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(721)
US-10-527-771-1

Query Match          35.8%; Score 273.4; DB 8; Length 828;
Best Local Similarity 63.6%; Pctd. No. 6,4e-71;
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2

QY      38 CTCCTGCGCCCTGTTCTCTATGCGCGAAGCAGGCTTTGTTGTCGGAATAGTCTAAAGCCAA 97
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      44 CTGGTCTGCTGTCATCAATATACCGAAGCAGGTTTTCCTGCCAGCAATCTTAACCA 103
QY      98 AGTGACAGCGCCGAGGAGCATTTTCTCGATTTTCACAATGATGTTGTCGAAATATAGCA 157
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      104 ACTGATGAGGCAAGMAAAATCTTCCTCGATTTTCACAATCAAGTTGCGGTGATATAGCA 163
QY      158 CTTGAAATGCTTTGATTAACCTGACACGTAATATGACAGCGGTCA---TTCTGCTCCA 214

```

Db 164 GGTGCAAGCCGGTGTCTCAACCTCACCGGAGCTGTTCARATGGCAATATGTTCTGGTCCA 223

Qy 215 GCTCAGAACATATACAAAGTGAAGTGGATTTGCAACTTGGAAAGATGACGACACACAG 274

Db 224 GCTAAGAACATATACAAATGGAATGGGATCTGGAACTGGAAAGCAAAAGCAAGCAATG 283

Qy 275 ATTGGCCATGCAATGATGCCCTTACCGATTAATACGACCTGGCTCAAAATATCGCTAGAC 334

Db 284 ATTGGCCATGCACTACGCCCTCTGGCAATAGACAGATATTCACAAAATCTCGCTCAC 343

Qy 335 TGGCTGTACTTCAAAAGACATGTAAGAAAGACAACTTTCGACAAAGATATGTTGG 394

Db 344 TGGCTGTACTTTCAAAACAGCTCAGGAATMTGAAGTGTGACGCAACGCCCTGGTCTTGG 403

Qy 395 GTGAGGCGATGCTGGGATTTATGAAAGGACAGAACTTGCACAAATTTGCTAACAGTGG 454

Db 404 GTACCGGACATACGAAATCTTCACTGTATACAGAGTAACTTTATTAATGGCA 463

Qy 455 GCTGAACCTCTAGCAACATTTGCAAACTATTAGAAACGAAAGTTGAGTGTGCCATAG 514

Db 464 ATTAGACCACTATCCACATTCGAACTGGCAAAACCTAAAGTTGATGTGCTCAAAA 523

Qy 515 ATCTG-----CCCGCTCAGCAAAACATGTGATATCTGCGTGTATGGAAGCCCAAA 568

Db 524 GTGTGCAAAATTCCTCCACCGGACAAATATGTGTGTGTCTTGGCTTATGGCGGCGAAGTA 583

Qy 569 CTTGACACCGAAGCAAGTTATCTTGCGAGGAAGAAAGCTTGTGTGTGCAAGCGTCGTCCA 628

Db 584 CTCCAAGATPACGAAGTTGTTATGGGACAAAGGACCAACCTTGCATGTGCAAAGTCTTATCC 643

Qy 629 GATTCAATCTGCTGCGCAAACTGTGTGACACGAGATGCTGCGAGATGTTCCGACACAG 688

Db 644 AACTCGTTGTGCTGCAACAATCTGTGTGACCAATATAGCTGCTGCGACACTTTCGACAG 703

Qy 689 TGTGTGCGGTGCGCATGAAGCAAAAGAAATTTGTAGTCAACC 731

Db 704 CCTGTGAATGCACTTGAAGCAAAAGGCGTTGTGTGATGTCCC 746

```

RESULT 3
US-11-121-086-42
Sequence 42, Application US/11121086
Publication No. US20050266459A1
GENERAL INFORMATION:
APPLICANT: FOULSEN, TIM S
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138, 6000-00000
CURRENT APPLICATION NUMBER: US/11/121, 086
CURRENT FILING DATE: 2005-05-04
PRIOR APPLICATION NUMBER: 60/567, 570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 42
LENGTH: 142303
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-42

Query Match          4.8%; Score 36.6; DB 12; Length 142303;
Best Local Similarity 56.1%; Pred. No. 13;
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

174 TAAACTGACAGTAAATGACGACGCGTCATTCTTGTCACAGTCAGAACTGTACAAAG 233
|||||
104467 TATCTGAACAAATTAAAGCAATTATTATGCGCTCTTATCTGTCTCCTATTAAG 104526

234 TGCATCGGATATGCAACTTGGAAAGTAGACGACACAAAGATGGCCATGCAATGATC 293
|||||
104527 TGCATCGGATATGCAACTTGGAAAGTAGACGACACAAAGATGGCCATGCAATGATC 104586

```

QY 294 CCC 296  
DB 104587 GCC 104589

RESULT 4  
US-09-925-065A-501602

Sequence 501602, Application US/09925065A  
Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 501602

LENGTH: 591

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-501602

Query Match 4.6%; Score 34.8; DB 6; Length 591;

Best Local Similarity 56.2%; Pred. No. 4;

Matches 63; Conservative 1; Mismatches 48; Indels 0; Gaps 0;

QY 322 AATATCGTAGATGCTGCTACTTCAAGACAGTGAAGAGACAGTTCGCAACACT 381

DB 215 AATATTAAGAGATGCTGCTATGCTATGCTATGAGATGAAGTAACTATGCTCAGAGAGA 274

QY 382 ATCGTGATTTGGTGAGCGCATCGCTGGATTGAAGAGCAGCAAACTT 433

DB 275 AAGAGATATGCTGTGTGATGACCTGAGTTAAGAGAGGCGCTTCCTT 326

RESULT 5

US-09-925-065A-611636/c

Sequence 611636, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 611636

LENGTH: 529

TYPE: DNA

ORGANISM: Homo sapiens  
US-09-925-065A-611636

Query Match 4.4%; Score 33.4; DB 6; Length 529;

Best Local Similarity 65.3%; Pred. No. 10;

Matches 49; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 124 CCAATTTTCAGATGATGTTGCTGCAAAATATGACACTTGGAATGTTGATTAACAGAC 183

DB 393 CAAATTTAAATTTATTTATATCATTAATTTGACTGGGAAGGGTTTTTAAATGAGAC 334

QY 184 AGTAATGCAAGCGC 198

DB 333 CTGAAGCATATGCG 319

RESULT 6  
US-09-925-065A-524748/c

Sequence 524748, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24

PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20

PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/289,846

NUMBER OF SEQ ID NOS: 957086

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 524748

LENGTH: 538

TYPE: DNA

ORGANISM: Homo sapiens

US-09-925-065A-524748

Query Match 4.4%; Score 33.4; DB 6; Length 538;

Best Local Similarity 51.7%; Pred. No. 10;

Matches 76; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 342 ACTTGAAGACAGTGAAGAGACAGTTCGCAACAGTATCGGTATTTGGGTGAGCG 401

DB 159 AGTGCAAGCCAGTGAATTAACAGCAAGTGGCAAGGAATGACATGCTCCGGTGAGCA 100

QY 402 CATCGCTGGATTTATGAAGGACGAACTGACCAATTTGCTAACGAGTGGCTGAAC 461

DB 99 GACTACTGCAATGAAAAGAGGACATCTTGTGCTTGAAGAAAAGAACAGCAAGT 40

QY 462 CTTAGCAACATTTGCAAACTATAGAA 488

DB 39 CCTCACAACATCTGATGCTGCCGA 13

RESULT 7  
US-09-925-065A-625826/c

Sequence 625826, Application US/09925065A

Publication No. US20040181048A1

GENERAL INFORMATION:

APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single

FILE REFERENCE: 108827.135

CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2001-08-08

Query Match	4.4%	Score 33.4;	DB 6;	Length 570;
Best Local Similarity	48.2%	Pred. No. 10;		
Matches 94;	Conservative 0;	Mismatches 101;	Indels 0;	Gaps 0
QY	468	CAACACTTGCAGAACTATGAAACCGGAAGGTGSGATGTGCCATTAAGATCTGCCCGCTC	527	

```

Query Match Similarity      4.4%; Score 33.4; DB 8; Length 149382;
Best Local Similarity      55.7%; Pred. No. 1.1e+02;
Matches      64; Conservative      0; Mismatches      51; Indels      0; Gaps      0;

      Oy      82 GAATAGCTTAAGCCAAAGTAGACAGCGCAGACGATTTTCCTCGATTTTCACAAATATGT 141
      |||||
      Db      24006 GAATGCGCTTAAGCTTAATATCATTTTAAATTAACCTGATATTAAAGTTTAAAGACCA 23947

      Oy      142 TCGTCGAATATTAGACCTTGGAATGGTTTGATTAACCTGACAGTAATGACAGC 196
      |||||
      Db      23946 TGATCATTATAGAACTAGAAATGTTTCAGAGAGCAATTAATTTATACAGAC 23892

RESULT 10
US-09-925-065A-301865
Sequence 301865, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925.065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 036
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846

```

; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 301865  
 ; LENGTH: 574  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-925-065A-301865

Query Match 4.3%; Score 33.2; DB 6; Length 574;  
 Best Local Similarity 52.1%; Pred. No. 12;  
 Matches 74; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 271 ACAGATGGCCATGATGATCCCTACGATTAATACAGCTGCTCAAAATATCCG 330  
 DB 221 ACTGATATCTGTAGGCAATATATACATGATTAATTTGTATCTTAACATATCTA 280  
 QY 331 TAGATGGCTGTACTTCAAGACAGTGAAGAGACAGTTTGCAACAGTATCTGTGA 390  
 DB 281 AACATCAATGATGATGATGATTAAGTAAATGATTAATAAATGATTAACCTGTA 340  
 QY 391 TTGGGTAGAGGAGATGCTGCGGA 412  
 DB 341 TAGGCAAGTACCATGATGGA 362

RESULT 11

; US-09-925-065A-625825/C  
 ; Sequence 625825, Application US/09925065A  
 ; Publication No. US20040181048A1  
 ; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single  
 ; Nucleotide Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.135  
 ; CURRENT APPLICATION NUMBER: US/09/925,065A  
 ; PRIOR FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: US 60/243,096  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 60/252,147  
 ; PRIOR FILING DATE: 2000-11-20  
 ; PRIOR APPLICATION NUMBER: US 60/250,092  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/261,766  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/289,846  
 ; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 625825  
 ; LENGTH: 570  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-925-065A-625825

Query Match 4.3%; Score 33; DB 6; Length 570;  
 Best Local Similarity 47.7%; Pred. No. 14;  
 Matches 93; Conservative 1; Mismatches 101; Indels 0; Gaps 0;

QY 468 CAACATTCGCAATATAGAAACGAAAGTTGATGTCCTCCATAGATCTCCCGCTC 527  
 DB 252 CTAGACACTTGAAGACGAGGACGAGCATCTTGAGTGAAGAGCTCGAGACAGCC 193  
 QY 538 AGCAAAACATGTATATCTCGGTATGAGAAAGCCCAACTTGACCGAAAGATTGA 587  
 DB 192 TGGTCAACATGATGATCTCTCTCTACTTAAAGTAAAGTTAGCTGGAATAGCTT 133  
 QY 588 TCTGAGAGAGAAAGAGCTTGTGTGAGAGCTGTGCTCAATTCATCTGCTGACA 647  
 DB 132 GAACCCGAGAGCGAGAGTGAAGTGAAGTGAATGTCACATGATTCACGCCAGGC 73  
 QY 648 AACTGTGACACGC 662

DB 72 AACAGATGAGACAC 58

RESULT 12

; US-09-925-065A-193665/C  
 ; Sequence 193665, Application US/09925065A  
 ; Publication No. US20040181048A1  
 ; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single  
 ; Nucleotide Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.135  
 ; CURRENT APPLICATION NUMBER: US/09/925,065A  
 ; PRIOR FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: US 60/243,096  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 60/252,147  
 ; PRIOR FILING DATE: 2000-11-20  
 ; PRIOR APPLICATION NUMBER: US 60/250,092  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/261,766  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/289,846  
 ; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 193665  
 ; LENGTH: 576  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-925-065A-193665

Query Match 4.3%; Score 32.6; DB 6; Length 576;  
 Best Local Similarity 49.7%; Pred. No. 18;  
 Matches 83; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 349 AGACAGTGAAGAGACAGTCTGCAACATATGCTGTGATGGTGAAGGCAATCGCT 408  
 DB 253 ACGACTACCCGAAACATCTGACACAGTCCCTAGTATTTGTTATCATCTGACA 193  
 QY 409 GGAATTTAGAAAGCAGCAACCTGACCAATTTGCTAACAGTGGGCTGAACCTTAC 468  
 DB 192 CTGATATCTAAATTCACCATCAACATGAGATGATATCACCTCTGAGTTTAC 133  
 QY 469 AACATTCGCAATATAGAAACGAAAGTTGATGTCCTCCATAGATCTCCCGCTC 515  
 DB 132 AAAATCTACTATATGTTAGTTGAGTGGTATCCCTAGGAGA 86

RESULT 13

; US-09-925-065A-745205  
 ; Sequence 745205, Application US/09925065A  
 ; Publication No. US20040181048A1  
 ; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single  
 ; Nucleotide Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.135  
 ; CURRENT APPLICATION NUMBER: US/09/925,065A  
 ; PRIOR FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: US 60/243,096  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 60/252,147  
 ; PRIOR FILING DATE: 2000-11-20  
 ; PRIOR APPLICATION NUMBER: US 60/250,092  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/261,766  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/289,846  
 ; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 745205

```
LENGTH: 612
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 501
OTHER INFORMATION: n = A,T,C or G
US-09-925-065A-745205
```

```
Query Match 4.3%; Score 32.6; DB 6; Length 612;
Best Local Similarity 55.9%; Pred. No. 18;
Matches 62; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
```

```
QY 118 TTTCCTGATTTTCAATGATGTCGCAATATACACTTGAATGTTGATTA 177
DB 189 TTTTCTACTTACAAACTATCATATATATATATATATATATATATATAT 248
QY 178 CTGACATAATGACGCGGCTGATTCCTGCTCAGCTCAGACATGTA 228
DB 249 ATGAGCTGTATTTAGGCTGCTTCACTGCTTTGTAACCTTATCATTA 299
```

RESULT 14

```
US-09-925-065A-189082/c
Sequence 189082, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 189082
LENGTH: 559
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-189082
```

```
Query Match 4.2%; Score 32.4; DB 6; Length 559;
Best Local Similarity 51.4%; Pred. No. 20;
Matches 72; Conservative 1; Mismatches 67; Indels 0; Gaps 0;
```

```
QY 339 TGTACTTCAAGACACTGAGAGAGACAGTCTGCAACAAGTATGTGTATGGGTGA 398
DB 535 TTTAATTAATATATCTGTAACATACAAACAGTTACCAAACTGTGTCTATCTAGA 476
QY 399 GCGCATCGCTGGGATTTATGAAAGCAGCAAACTTGACCAATTTGTAACCGATGGGCTG 458
DB 475 GTGACGACAGGAAATTAACATGATTAACCTTACCAATGTAAGTAAGATCTG 416
QY 459 AACCTGACGAACATTTGA 478
DB 415 AACCAATTTGCCAAAATCCA 396
```

```
RESULT 15
US-09-925-065A-250352/c
Sequence 250352, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
```

```
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT FILING DATE: US/09/925,065A
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 250352
LENGTH: 591
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-250352
```

```
Query Match 4.2%; Score 32.4; DB 6; Length 591;
Best Local Similarity 46.9%; Pred. No. 21;
Matches 169; Conservative 0; Mismatches 186; Indels 5; Gaps 2;
```

```
QY 6 CTGCTATGTCGGGCGCTGTTGATGTCGTCCTTCCTGCGCCCTTCCTATGCCAG 65
DB 581 CAGCTACTTCTTACCAATGGGCATTAAGATTTGATTTACCTCTTATGTTGCTTA 522
QY 66 CAGCTTTGTTGTCGCAATAGCTTAAGCCAAAGTGAAGCGGAGGAGATTTCTCG 125
DB 521 CAGATACCAACCAACCAATTTATTAAGAAACTGACAGATTAATTAATTTGCTTAC 462
QY 126 ATTTTCAATGATGTTGCGAAATATAGCACTTGAATGTTGATTAACCTGACAG 185
DB 461 AATTTAATGATTTATTTATGCAAGAGGATCTTTTGAAGTGAAGGATTAAGTGAAG 402
QY 186 TAAATGACAGCGGCTCATTTGTCCTGACCTGACAAACATGTACAAAGTGAAGT 245
DB 401 TGAAGCCAGGGGCTGGGCTTAT--TCCAGACTTGATGACTGTAACTTTATAGTAT 345
QY 246 GCAACTTGAAGAGTACAGACAAACAGAT--TGGCCATGACATGATCCCTTACCAT 303
DB 344 TCAACAGAAATATGATCATCAAAACATGAGATGACCACTAGAGACTGTTCTTTGGCA 285
QY 304 AATATCAGAGCTGCTCAAAATATGCTAGATGCTGTACTTCAAAAGACGTGAAGAGA 363
DB 284 AGCAGCTGATGATGATTAATTTGAGCAACTACCTGCTTAAAGATGCTTAAGAAAA 225
```

```
Search completed: March 4, 2006, 07:35:38
Job time : 336 secs
```



Yr	Sequence	Position
Qy	1 GAGAACTGCTAAGTCGGAGCGGCTGTGTAGTGTCTCTGGACCCGTGTCTTAAGC	60
Qy	2 GAGAACTGCTAAGTCGGAGCGGCTGTGTAGTGTCTCTGGACCCGTGTCTTAAGC	60
Db	1 GAGAACTGCTAAGTCGGAGCGGCTGTGTAGTGTCTCTGGACCCGTGTCTTAAGC	60
Qy	61 CGAAGCAGGCTTTGTGTTCGGAATAGCTAAGCCAAAGTACAGCGCGAGCGAGATTTT	120
Db	61 CGAAGCAGGCTTTGTGTTCGGAATAGCTAAGCCAAAGTACAGCGCGAGCGAGATTTT	120
Qy	121 CCGCATTTTCAACATAGATGTTCCGCAATATATAGCACTTGGAAATGATTTGATTAATCTG	180
Db	121 CCGCATTTTCAACATAGATGTTCCGCAATATATAGCACTTGGAAATGATTTGATTAATCTG	180
Qy	181 GACAGTAAATGCAAGCGCGGTCAATCTTGGTTCAGGCTCAGAACTGTACAAAGTGAATG	240
Db	181 GACAGTAAATGCAAGCGCGGTCAATCTTGGTTCAGGCTCAGAACTGTACAAAGTGAATG	240
Qy	241 GGAATTCACACTGGAGAAAGTATAGCAGCAACAAGATTTGCGCAATGATATCCCTTAAC	300
Db	241 GGAATTCACACTGGAGAAAGTATAGCAGCAACAAGATTTGCGCAATGATATCCCTTAAC	300
Yr	301 GATTAATACAGACTGGCTCAAAATATGCTAAGATGCTGTACTTCAAGACATGTAGAGA	360

```

Db 301 GATTAATACAGACCTGGCTCAAAAATATGCTAGATGCTGTACTTCAAAAGACATGAAAGA 360
Qy 361 AGAGACAGTTCTGCAACAGATATCTGTGATTTGGGTGAGCCGATGCTGGGATTTATGAA 420
Db 361 AGAGACAGTTCTGCAACAGATATCTGTGATTTGGGTGAGCCGATGCTGGGATTTATGAA 420
Qy 421 AGGCGCAAACTTGGCAATTTGCTAACAGTGGGCTGAACCTTAGAGCAATTTGCAAA 480
Db 421 AGGCGCAAACTTGGCAATTTGCTAACAGTGGGCTGAACCTTAGAGCAATTTGCAAA 480
Qy 481 CTATGAAACCGAAAGTTGATGTGATGCTGATTAAGATCTGCCCCGCTGACAAACATGAT 540
Db 481 CTATGAAACCGAAAGTTGATGTGATGCTGATTAAGATCTGCCCCGCTGACAAACATGAT 540
Qy 541 AGTATCTGCGTGTATGAAAGCCCAAACTTGACCGAAGCAAGTTATCTGGCAGAAAG 600
Db 541 AGTATCTGCGTGTATGAAAGCCCAAACTTGACCGAAGCAAGTTATCTGGCAGAAAG 600
Qy 601 AAAGCTTGTGTGTGCGAGAGCTGCTGCTGCAATTCATTCGTGCGCAACCTGTGTGACAC 660
Db 601 AAAGCTTGTGTGTGCGAGAGCTGCTGCTGCAATTCATTCGTGCGCAACCTGTGTGACAC 660
Qy 661 GCGAGATGCTGCGAGATGTTGCGCAACAGTGTGCGGCTGCGCATGAAAGCAAAAGAAAT 720
Db 661 GCGAGATGCTGCGAGATGTTGCGCAACAGTGTGCGGCTGCGCATGAAAGCAAAAGAAAT 720
Qy 721 GGTAGTCAACCCGCAATTAATAATTCATGCAAAAAA 763
Db 721 GGTAGTCAACCCGCAATTAATAATTCATGCAAAAAA 763

```

```

RESULT 2
US-10-243-319C-7
; Sequence 7, Application US/10243319C
; Publication No. US20040052817A1
; GENERAL INFORMATION:
; APPLICANT: Universteit Gent
; TITLE OF INVENTION: Osteotegia vaccine
; FILE REFERENCE: Unigent Osteotegia
; CURRENT APPLICATION NUMBER: US/10/243,319C
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent version 3.2
; SEQ ID NO 7
; LENGTH: 828
; TYPE: DNA
; ORGANISM: Osteotegia osteotagi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(721)
US-10-243-319C-7

```

```

Query Match 35.8%; Score 273.4; DB 7; Length 828;
Best Local Similarity 63.6%; Pred. No. 1.9e-78;
Matches 447; Conservative 3; Mismatches 244; Indels 9; Gaps 2;
Qy 38 CTCCTGGCCCTGTTCTCTCTATGCGGAAGCAGCGCTTTGTTGTCGATAGTCTTAAGCAA 97
Db 44 CTCGTGTGTGTGATCAATCAATACGAAAGAGGTTTTGTCGCGCAAGATCTTAACCAA 103
Qy 98 AGTGAACGCGCGAGAGATTTTCTGATTTTCAATGATGTTGTCGAATATATGCA 157
Db 104 ACTGATGAGGCAAGAAATCTTCTGATTTTCAATCAATCAATGTCGCCCTGTATATAGCA 163
Qy 158 CTTGGAATGTTGATTAATCTGACATTAATGACAGCGGTC---TTCTTGATCA 214
Db 164 GGTGCAAGCCCGTTGCTCAACCTCAACGAGCTGTTCARATGCAAAATGTTCTGGTCA 223
Qy 215 GCTCAAGAACATGTAAGATGAGTGGGATTTGCAATTTGCAAGAAAGTACAGCAACAG 274
Db 224 GCTTAAGAACATGTAAGATGAGTGGGATTTGCAATTTGCAAGAAAGTACAGCAACAG 283

```

```

Qy 275 ATTGCGCATGATGATGATCCCTTACCGATTAATACAGCCTGGCTCAAAATATGCTAGA 334
Db 284 ATTGGCCATGACCTAGCCCTCTGCCCAATAGACAGATATTCACAAATCTGCTCAR 343
Qy 335 TGGGTGATCTTAAAGACAGTGAAGAGAGACAGTTCTGCAACAAATGATGCTGATTTGG 394
Db 344 TGGGTGATCTTAAAGACAGTGAAGAGAGAGATGATGATGATGATGATGATGATGATGAT 403
Qy 395 GTGAGCGCATGCTGGGATTTATGAAAGCAAGCAACTTGAACAAATTTGCTAACAGTGG 454
Db 404 GTAACCCGATACATGAGAAATCTTCAACCTGATACAGAGTAACTATTATTAATGACGAA 463
Qy 455 GCTGAACCTTGAAGAACTTGAACATTAATGAACCGAAAGTTGATGATGATGATGATGAT 514
Db 464 ATTAGACCATATCAATGACATGCGAAGCACTTAAGTTGATGATGATGATGATGATGATGAT 523
Qy 515 ATCTG-----CCCCGCTGACGAAACAGTGTATGATGATGATGATGATGATGATGATGAT 568
Db 524 GTGTGCAAAATTCCTCCACCGGACAAATATGATGATGATGATGATGATGATGATGATGAT 583
Qy 569 CTGCAACCGAAGAAATTTATCTGCGAGAGAAAGGCTGTGTGTGCGACGCTGTTCA 628
Db 584 CTCCAAATTAACGAAGTTGATGAGGACAGGACCAACTTGATGATGATGATGATGATGATGAT 643
Qy 629 GATTCATTTCTGCTGCGACCACTGTGTGACACGCGAGATGCTGCGAGTTTGGCCACAG 688
Db 644 AACTGTTCTGCTGCAACATCTGTGTGACCAATATGCTGCGACCACTTGTGCAACAG 703
Qy 689 TGTGCGGCTGCGCATGAAAGGAAATTTGATGACCC 731
Db 704 CTTGTAATATGACTGTGAAGGAAAGCGCTGTGTGATGCTCC 746

```

```

RESULT 3
US-10-388-838-84
; Sequence 84, Application US/10388838
; Publication No. US20040180344A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer
; FILE REFERENCE: 529452001600
; CURRENT APPLICATION NUMBER: US/10/388,838
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 260803
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(260803)
; OTHER INFORMATION: n = A,T,C or G
US-10-388-838-84

```

```

Query Match 4.7%; Score 36; DB 8; Length 260803;
Best Local Similarity 45.7%; Pred. No. 34;
Matches 126; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
Qy 103 CAGCGGAGGAGATTTTCTGATTTTCAATGATGTTGTCGAATATATGCACTTGG 162
Db 259473 CTGGGAAAGATTAATATCCCACTTATCTGAATGATGATGATGATGATGATGATGAT 259532
Qy 163 AAATGTTGATTAATCTGACATTAATGACAGCGGTCATTTCTGTGCTGACGCTGAA 222
Db 259533 GATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 259592
Qy 223 CATGTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
Db 259593 AAATTTGGGGGTGCTTTGTTAAGCAGCTAGAGAGATTAACCTGATGATGATGATGAT 259652
Qy 283 ATGCAATGATCCCTACGATTAATATACAGCCTGCTCAAAATATGCTAATGCTGTA 342

```

DB 259653 ATGCATCTGAGATTACAAATCCAGTATGATTTTCACTTCTGATTGA 259712  
QY 343 CTTCAAGACAGTGAAGAGACAGTTCTGACACA 378  
DB 259713 CCACATCCATTAGAGGACAGAAAATTCACACACA 259748

## RESULT 4

US-09-925-065A-501602  
; Sequence 501602, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 501602  
; LENGTH: 591  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-501602

Query Match  
Best Local Similarity 4.6%; Score 34.8; DB 4; Length 591;  
Matches 63; Conservative 1; Mismatches 48; Indels 0; Gaps 0;

QY 322 AATATCGCTAGATGCTGTACTTCAAGACAGTGAAGACAGTTCTGCACACAGT 381  
DB 215 AATATTAAGAGATGCTGTATGCTATGATGAATGAACAGAAAATCTGACAGAGA 274  
QY 382 ATCTGTGATTTGGGTGAGCGCATCGCTGGATTATGAAGCAGCAAACTT 433  
DB 275 AAGGAGTATATGCTGTGTAATGACCCGTGATTAGAGAGGCGCTTCCTT 326

## RESULT 5

US-10-741-601-5625/c  
; Sequence 5625, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5625  
; LENGTH: 16977  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-601-5625

Query Match  
Best Local Similarity 4.5%; Score 34.6; DB 7; Length 16977;  
Matches 100; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 555 ATGAGAGCCCAAACTTGACACCGAAGATTATCTGGAGAGAGAAAGCTTGTTGT 614  
DB 11730 ATGAGAGAAACCCGTGTCTTACTTAAACAAATATAGCTGGGCGTGGTGTCT 11671  
QY 615 GCGAGCTGTGTCAGATTATTTCTGCTGCAACAACCTGTGTGACAGCGAGTCTGCA 674  
DB 11670 GTAAATCCAGCTACTCAGAGAGCTGAGGAGAAATCTGAACTGAAGAGGAGAG 11611  
QY 675 GTGTTGCCACAGATGTTGCGGCTGCATGAAGAGCAAAATGTGTACCCCGA 734  
DB 11610 TTGTGTGAACCGAATGACGCGTTTCCTTCACTGCTGGGCAACAAAGCAAACTCCG 11551  
QY 735 ATAAATATTTCATGCAAAAAA 763  
DB 11550 CTCAAAAA 11522

## RESULT 6

US-10-087-192-826/c  
; Sequence 826, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; TITLE OF INVENTION: CANCER  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; CURRENT FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 826  
; LENGTH: 194883  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-087-192-826

Query Match  
Best Local Similarity 4.5%; Score 34.6; DB 5; Length 194883;  
Matches 82; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 150 ATATGACATTGAAATGTTGATTAATCTGACAGTAATGACAGCGGTCAATCTTG 209  
DB 79227 AGATGACAGTGCCTGTATGCTCCAGCTACTCAGAGAGAGAGTGAAGAGATCATCG 79168  
QY 210 GTCCAGCTCAGAAACATGTAACAAGTGAAGTGGATTGCACTTGAAGAAGTAGCAGC 269  
DB 79167 GAGCCAGAGATTCAAGGCTGAAGTGGCTATGATGTCATGTGTAAATGACCATGTAC 79108  
QY 270 AACAGATTGCGCCATGATGATCCCTTACCGATTAATACC 310  
DB 79107 TCCAGCTGGCAACACAGCGGAACCTTATCTAAGAAAATC 79067

## RESULT 7

US-10-367-094-122/c  
; Sequence 122, Application US/10367094  
; Publication No. US20040170982A1  
; GENERAL INFORMATION:  
; APPLICANT: David W. Morris  
; APPLICANT: Marc Malandro  
; TITLE OF INVENTION: Novel Therapeutic Targets in Cancer  
; FILE REFERENCE: 529452001500  
; CURRENT APPLICATION NUMBER: US/10/367,094  
; CURRENT FILING DATE: 2003-02-14  
; NUMBER OF SEQ ID NOS: 203  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122  
; LENGTH: 421609

;; TYPE: DNA  
;; ORGANISM: Mus musculus  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: (1) ... (421609)  
;; OTHER INFORMATION: n = A,T,C or G  
US-10-367-094-122

Query Match 4.5%; Score 34.6; DB 7; Length 421609;  
Best Local Similarity 50.3%; Pred. No. 1.3e+02;  
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 509 CATGAAGTCTGCCCCGCTCAGCAAAACATGTAGTATCTGCTGATGGAAGCCCAAA 568  
DB 65832 CTTAGACATGGGCGCCGCTGCGATGAGGTTACCTTATGACGGGTAAAGCTGTGACATT 65773  
QY 569 CTGACACCAAGCAAGTTATCTGCGAGAGAAAGCTTGTGTGCGACGCTGTGCA 628  
DB 65772 AGAGAAACGAACTTAAACAGAGCCAGCGGATGGCATTAATTCACAGGCTTAGTCCA 65713  
QY 629 GATTGATTCTGCTGCGACAACTGTGTGACACGCGAGATGCTGCGAGTG 677  
DB 65712 GCCTATTATTATTAACAAACAGAGGCGAGATGTGGAAAGCCACATGTG 65664

RESULT 8  
US-10-363-345A-10699/C  
; Sequence 10699, Application US/10363345A  
; Publication No. US20040234960A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Olek  
; APPLICANT: Christian Piepenbrock  
; APPLICANT: Kurt Berlin  
; TITLE OF INVENTION: Method for determining the degree of methylation of defined  
; FILE REFERENCE: E01/1227  
; CURRENT APPLICATION NUMBER: US/10/363,345A  
; CURRENT FILING DATE: 2003-03-03  
; NUMBER OF SEQ ID NOS: 40712  
; SEQ ID NO 10699  
; LENGTH: 1003  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; OTHER INFORMATION: CpG-island No: 10699  
US-10-363-345A-10699

Query Match 4.5%; Score 34.4; DB 8; Length 1003;  
Best Local Similarity 52.0%; Pred. No. 6.2; 71; Indels 0; Gaps 0;  
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTGCTGCAATTCATTCTGTCGCAACCTGTGTGACACCGAGATGCTGCGAG 675  
DB 986 CGGCTCCGTCGCAACCCCTACCTACGACGACGACTATAAATACTCAAA 927  
QY 676 TGTTCGCCACCAAGTGTGCGCGCTGCGCATGAGCAAAAGAAATGTAATCAACCCGAA 735  
DB 926 CGCTGCGCCGGAATTAAACAGAGAAAAAACAATATAAAGCAAAACAAATCGAA 867  
QY 736 TAAATATTCATGCAAAAAA 763  
DB 866 TACATTAACCCGAAAAAAGCAAAA 839

RESULT 9  
US-10-363-345A-10700  
; Sequence 10700, Application US/10363345A  
; Publication No. US20040234960A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Olek  
; APPLICANT: Christian Piepenbrock  
; APPLICANT: Kurt Berlin

;; TITLE OF INVENTION: Method for determining the degree of methylation of defined  
;; FILE REFERENCE: E01/1227  
;; CURRENT APPLICATION NUMBER: US/10/363,345A  
;; CURRENT FILING DATE: 2003-03-03  
;; NUMBER OF SEQ ID NOS: 40712  
;; SEQ ID NO 10700  
;; LENGTH: 1003  
;; TYPE: DNA  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
;; OTHER INFORMATION: CpG-island No: 10700  
US-10-363-345A-10700

Query Match 4.5%; Score 34.4; DB 8; Length 1003;  
Best Local Similarity 52.0%; Pred. No. 6.2; 71; Indels 0; Gaps 0;  
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTGCTGCAATTCATTCTGTCGCAACCTGTGTGACACCGAGATGCTGCGAG 675  
DB 18 CGGCTCCGTCGCAACCCCTACCTACGACGACGACTATAAATACTCAAA 77  
QY 676 TGTTCGCCACCAAGTGTGCGCGCTGCGCATGAGCAAAAGAAATGTAATCAACCCGAA 735  
DB 78 CGCTGCGCCGGAATTAAACAGAGAAAAAACAATATAAAGCAAAACAAATCGAA 137  
QY 736 TAAATATTCATGCAAAAAA 763  
DB 138 TACATTAACCCGAAAAAAGCAAAA 165

RESULT 10  
US-10-363-483A-10699/C  
; Sequence 10699, Application US/10363483A  
; Publication No. US20050064401A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Olek  
; APPLICANT: Christian Piepenbrock  
; APPLICANT: Kurt Berlin  
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain  
; FILE REFERENCE: 82011  
; CURRENT APPLICATION NUMBER: US/10/363,483A  
; CURRENT FILING DATE: 2003-03-03  
; NUMBER OF SEQ ID NOS: 40712  
; SEQ ID NO 10699  
; LENGTH: 1003  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; OTHER INFORMATION: CpG-island No: 10699  
US-10-363-483A-10699

Query Match 4.5%; Score 34.4; DB 9; Length 1003;  
Best Local Similarity 52.0%; Pred. No. 6.2; 71; Indels 0; Gaps 0;  
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTGCTGCAATTCATTCTGTCGCAACCTGTGTGACACCGAGATGCTGCGAG 675  
DB 986 CGGCTCCGTCGCAACCCCTACCTACGACGACGACTATAAATACTCAAA 927  
QY 676 TGTTCGCCACCAAGTGTGCGCGCTGCGCATGAGCAAAAGAAATGTAATCAACCCGAA 735  
DB 926 CGCTGCGCCGGAATTAAACAGAGAAAAAACAATATAAAGCAAAACAAATCGAA 867  
QY 736 TAAATATTCATGCAAAAAA 763  
DB 866 TACATTAACCCGAAAAAAGCAAAA 839

RESULT 11

US-10-363-483A-10700  
; Sequence 10700, Application US/10363483A  
; Publication No. US20050064401A1  
; GENERAL INFORMATION:  
; APPLICANT: Alexander Olek  
; APPLICANT: Christian Plepenbrock  
; APPLICANT: Kurt Berlin  
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain  
; FILE REFERENCE: 82011  
; CURRENT APPLICATION NUMBER: US/10/363,483A  
; NUMBER OF SEQ ID NOS: 40712  
; SEQ ID NO 10700  
; LENGTH: 1003  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; OTHER INFORMATION: Cpg-Island No: 10700  
US-10-363-483A-10700

Query Match 4.5%; Score 34.4; DB 9; Length 1003;  
Best Local Similarity 52.0%; Pred. No. 6.2;  
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 616 CGACGCTGCTCCAGATTCTTGTGCTGCGACCACTGTGTGACGCGAGATGCTGCGAG 675  
DB 18 CGCGTCCCTCCGACACCTTACCTTACGCGACGACATTAACAATACTTAACCTCCAAA 77  
QY 676 TGTTCGCCACAGTGTGCGCTGCGCATGAGCGAAAGAAATTTGTAGTCACTCCGGA 735  
DB 78 CGCTGCGCCCGCACTTTAAGCAGCAAAAAAACCAATTAACGAAACAAATCCGA 137  
QY 736 TAAATATTGATGCAAAAAA 763  
DB 138 TACAAATAAACCCGAAAAAAGCAAAA 165

## RESULT 12

US-10-027-632-97924/c  
; Sequence 97924, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 97924  
; LENGTH: 2148  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-97924

Query Match 4.5%; Score 34.4; DB 5; Length 2148;

Best Local Similarity 53.8%; Pred. No. 9.3;  
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 120 TCCCTGATTTTCACATGATGTTGTCGAAATATGCACTTGAATGTTGATTAAC 179  
DB 765 TCCAGAAATTAACATCTTATTTCACTTAATAATTAATTAATTAATTAATTAATCA 706  
QY 180 GGACAGTAATGACAGCGCGTCACTTGTGTCAGCTCAGAAATGTAACAAGTGACT 239  
DB 705 GGTCTGTAATTAATTTTACCAATACCACTTGCCAGAACCAATGGGTTGTGAAGATGAT 646  
QY 240 GGAATGCACT 251  
DB 645 AAGACGCTCT 634

## RESULT 13

US-10-027-632-97924/c  
; Sequence 97924, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 97924  
; LENGTH: 2148  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-97924

Query Match 4.5%; Score 34.4; DB 6; Length 2148;  
Best Local Similarity 53.8%; Pred. No. 9.3;  
Matches 71; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 120 TCCCTGATTTTCACATGATGTTGTCGAAATATGCACTTGAATGTTGATTAAC 179  
DB 765 TCCAGAAATTAACATCTTATTTCACTTAATAATTAATTAATTAATTAATTAATCA 706  
QY 180 GGACAGTAATGACAGCGCGTCACTTGTGTCAGCTCAGAAATGTAACAAGTGACT 239  
DB 705 GGTCTGTAATTAATTTTACCAATACCACTTGCCAGAACCAATGGGTTGTGAAGATGAT 646  
QY 240 GGAATGCACT 251  
DB 645 AAGACGCTCT 634

## RESULT 14

US-10-437-963-51217/c  
; Sequence 51217, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2006, 07:13:05 ; Search time 3916 Seconds  
(without alignments)  
916.074 Million cell updates/sec

Title: US-10-527-771-9  
Perfect score: 763  
Sequence: 1 ggaactgctacgtcgcg...tcacgaaaaaaaaaaaaa 763

Scoring table: IDENTITY\_NIC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_ests:  
2: gb\_ests:  
3: gb\_ests:  
4: gb\_hic:  
5: gb\_ests:  
6: gb\_ests:  
7: gb\_ests:  
8: gb\_ests:  
9: gb\_gest:  
10: gb\_gest:  
11: gb\_gest:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	478.4	62.7	766	6	CB036994 Tc_ad2_28
2	476.8	62.5	757	6	CB038073 Tc_ad2_41
3	475.8	62.4	757	6	CB036437 Tc_ad2_21
4	474.2	62.1	760	6	CB038292 Tc_ad2_44
5	474.2	62.1	766	6	CB038112 Tc_ad2_42
6	473.6	62.1	756	6	CB036381 Tc_ad2_21
7	473.2	62.0	755	6	CB036975 Tc_ad2_28
8	472.6	61.9	751	6	CB038323 Tc_ad2_44
9	472.6	61.9	754	6	CB038204 Tc_ad2_43
10	470.2	61.6	771	6	CB038345 Tc_ad2_44
11	469.2	61.5	733	6	CB037619 Tc_ad2_36
12	469.2	61.5	763	6	CB037329 Tc_ad2_32
13	467.6	61.3	739	6	CB036489 Tc_ad2_22
14	466.2	61.1	800	6	CB037895 Tc_ad2_39
15	466.2	61.1	758	6	CB036762 Tc_ad2_25
16	464.6	60.9	721	3	BM052055 Tc_ad2_02
17	464.6	60.9	752	6	CB036281 Tc_ad2_19
18	464.6	60.9	757	6	CB036386 Tc_ad2_21
19	464.6	60.8	783	6	CB036875 Tc_ad2_27
20	464.2	60.8	736	6	CB037631 Tc_ad2_36
21	463	60.7	768	6	CB037578 Tc_ad2_35
22	463	60.7	778	6	CB038400 Tc_ad2_45

23	462.2	60.6	777	6	CB038358 Tc_ad2_44
24	462	60.6	783	6	CB038266 Tc_ad2_43
25	461.4	60.5	757	6	CB038393 Tc_ad2_45
26	459.2	60.2	704	6	CB037577 Tc_ad2_35
27	457.8	60.0	728	6	CB036543 Tc_ad2_23
28	457.4	59.9	711	6	CB038828 Tc_ad2_50
29	457	59.9	759	6	CB037859 Tc_ad2_39
30	456.2	59.8	709	6	CB036361 Tc_ad2_20
31	454.6	59.6	739	6	CB036890 Tc_ad2_27
32	454.4	59.6	704	6	CB038110 Tc_ad2_42
33	454.4	59.6	720	6	CB036228 Tc_ad2_19
34	454	59.5	767	6	CB036381 Tc_ad2_21
35	453.4	59.4	704	6	CB038648 Tc_ad2_48
36	453.2	59.4	756	6	CB038256 Tc_ad2_43
37	453	59.4	743	6	CB038839 Tc_ad2_50
38	452.8	59.3	735	6	CB036573 Tc_ad2_23
39	451.6	59.2	734	6	CB037846 Tc_ad2_38
40	451.2	59.1	728	6	CB037691 Tc_ad2_37
41	450.8	59.1	695	6	CB039176 Tc_ad2_54
42	450.2	59.0	703	6	CB036788 Tc_ad2_26
43	450.2	59.0	760	6	CB037386 Tc_ad2_33
44	449.4	58.9	700	6	CB038612 Tc_ad2_48
45	449.4	58.9	743	6	CB037372 Tc_ad2_33

## ALIGNMENTS

RESULT 1  
LOCUS CB036994 766 bp mRNA linear EST 15-JAN-2003  
DEFINITION Tc\_ad2\_28P03\_TEXPI Teladorsagia circumcincta adult library 2  
Teladorsagia circumcincta CDNA clone Tc\_ad2\_28P03\_5' similar to  
CAD56559 Ancylostoma-secreted protein-like prote Osteragia, mRNA  
sequence.

ACCESSION CB036994 GI:27756239  
VERSION CB036994  
KEYWORDS EST  
SOURCE Teladorsagia circumcincta  
ORGANISM Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongylidae; Haemonchidae; Oestertagiinae; Teladorsagia.

REFERENCE 1 (bases 1 to 766)  
Blaxter, M.L., Parkinson, J., Whitton, C., Dab, J., Guilianno, D.,

AUTHORS Hall, N., Quayle, W., and Barrell, B.  
TITLE Edinburgh University/Sanger Centre Nematode EST Project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk

The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).

PCR PRIMERs  
FORWARD: TEXPCRF1  
BACKWARD: T7PL

Plate: 28 row: F column: 03  
Seq primer: TEXP1  
High quality sequence stop: 519.

## FEATURES

source location/Qualifiers  
1..766  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="Tc\_ad2\_28P03"  
/sex="mixed"

ORIGIN

/dev stage="adult"  
/clone.lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTribEx2, site 1: BamHI, site 2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

Query Match 62.7%; Score 478.4; DB 6; Length 766;  
Best Local Similarity 79.8%; Pred. No. 1.1e-133;  
Matches 579; Conservative 0; Mismatches 141; Indels 6; Gaps 1;

19 GGCCTGTGAGTGGCTGTTCTCTGACCCTGTTCTCTATGCGGAAGAGGCTTTGTTG 78  
25 GGCAGTCAAGTGGCTGTTCTCTGACCCTGTTCTCTATGCGGAAGAGGCTTTGTTG 84  
79 TCCGAATAGTCTAACCCAAAGTGAACGCGGAGGAGATTTTCTCGATTTTCAATGA 138  
85 CCCAAACTCTAACGCAAACTGAACGCGGAGGAGATTTTCTCGATTTTCAATGA 144  
139 TGTGTGCGAAATATGACCTTGGAATGTTGATTAATCTGACAGTAAATGACAGCC 198  
145 TATTCGTGCGAATATGACCTTGGAATGTTGATTAATCTGACAGTAAATGACAGCC 198  
199 GGTCAATCTTGTCAGCTCAGAACATGTATACAAAGTGAAGTGGATTTGCAATGGA 258  
199 GGTGTCTTGTGTCAGCTCAGAACATGTATACAAAGTGAAGTGGATTTGCAATGGA 258  
259 AGTAGCAGCAACAGATTCGCGCATGCAATGATCCCTTACGATTAATACAGCTGGC 318  
259 AAAAGAGTCAACAGATTCGCGCATGCAATGATCCCTTACGATTAATACAGCTGGC 318  
319 TCAAAATATCGCTAGTGGCTGTACTTCAAAAGCATGGAAGAGACAGTCTGCAACA 378  
319 TCAAAATATCGCTAGTGGCTGTACTTCAAAAGCATGGAAGAGACAGTCTGCAACA 378  
379 AGTATCTGTGATGTTGGGTGAGGCAATCGCTGGGATTTATGAAGGCAAGAACTTGACA 438  
379 AGTATCTGTGATGTTGGGTGAGGCAATCGCTGGGATTTATGAAGGCAAGAACTTGACA 438  
439 ATTTGCTAACAGTGGCTGAACTCTTACCAAACTATGCAAACTATGAAACCGAAAGGT 498  
439 TTTTGTGTAACAGTGGCTGAACTCTTACCAAACTATGCAAACTATGAAACCGAAAGGT 498  
499 TGGATGTGCCATTAAGATTCGCCCGCTCAGCAAAACATGATGATTCCTGCGTGTATG 558  
499 TGGATGTGCCATTAAGATTCGCCCGCTCAGCAAAACATGATGATTCCTGCGTGTATG 558  
559 AAGCCCAAACTTGAACCGAAGATTTATCTGGCAGGAAGAAAGCTTGTGTGTGCA 618  
559 AAGCCCAAACTTGAACCGAAGATTTATCTGGCAGGAAGAAAGCTTGTGTGTGCA 618  
619 CGCTGTGCAAGTCAATTCCTGTCGCAACCTGTGTGACAGCGAGATGCTGCGAGTGT 678  
619 CTCTTATCCGAGATTCCTGTCGTCGCAACCTGTGTGACAGCGAGATGCTGCGAGTGT 678  
679 TCGCCACCAAGTGTGCGCGCTGCGCAATGAAAGGAAATTTGTAGTCAACCCCAATTA 738  
679 TCGCCACCAAGTGTGCGCGCTGCGCAATGAAAGGAAATTTGTAGTCAACCCCAATTA 738  
739 AATATT 744  
739 AATATT 744

RESULT 2  
CB038073 757 bp mRNA linear EST 15-JAN-2003  
LOCUS  
DEFINITION  
Teladorsagia circumcincta cDNA clone Teladorsagia circumcincta adults library 2  
CAD5659 Ancylostoma-secreted protein-like protease Ostertagia, mRNA  
sequence.  
CB038073

VERSION  
CB038073.1 GI:2757318  
EST

KEYWORDS  
Teladorsagia circumcincta

SOURCE  
Teladorsagia circumcincta

ORGANISM  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Strongylida;  
Trichostrongylidae; Haemonchidae; Ostertaginae; Teladorsagia.

REFERENCE  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guilianno, D.,  
1 (bases 1 to 757)  
Hall, N., Quayle, M., and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)

AUTHORS  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guilianno, D.,  
Hall, N., Quayle, M., and Barrell, B.

TITLE  
Edinburgh University/Sanger Centre Nematode EST Project

JOURNAL  
Unpublished (2000)

COMMENT  
Contact: Blaxter ML  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Main Road, Edinburgh, EH9  
3JF, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRP1  
BACKWARD: T7PL  
Plate: 41 row: B column: 11  
Seq primer: TEXP1  
High quality sequence start: 5  
High quality sequence stop: 519.  
Location/Qualifiers  
1..757  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="TC ad2 41E11"  
/sex="mixed"  
/dev stage="adult"  
/note="Vector: pTribEx2, site 1: BamHI, site 2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 62.5%; Score 476.8; DB 6; Length 757;  
Best Local Similarity 79.6%; Pred. No. 3.4e-133;  
Matches 578; Conservative 0; Mismatches 142; Indels 6; Gaps 1;

19 GGCCTGTGAGTGGCTGTTCTCTGACCCTGTTCTCTATGCGGAAGAGGCTTTGTTG 78  
25 GGCAGTCAAGTGGCTGTTCTCTGACCCTGTTCTCTATGCGGAAGAGGCTTTGTTG 84  
79 TCCGAATAGTCTAACCCAAAGTGAACGCGGAGGAGATTTTCTCGATTTTCAATGA 138  
85 CCCAAACTCTAACGCAAACTGAACGCGGAGGAGATTTTCTCGATTTTCAATGA 144  
139 TGTGTGCGAAATATGACCTTGGAATGTTGATTAATCTGACAGTAAATGACAGCC 198  
145 TATTCGTGCGAATATGACCTTGGAATGTTGATTAATCTGACAGTAAATGACAGCC 198  
199 GGTCAATCTTGTCAGCTCAGAACATGTATACAAAGTGAAGTGGATTTGCAATGGA 258  
199 GGTGTCTTGTGTCAGCTCAGAACATGTATACAAAGTGAAGTGGATTTGCAATGGA 258  
259 AGTAGCAGCAACAGATTCGCGCATGCAATGATCCCTTACGATTAATACAGCTGGC 318  
259 AAAAGAGTCAACAGATTCGCGCATGCAATGATCCCTTACGATTAATACAGCTGGC 318  
319 TCAAAATATCGCTAGTGGCTGTACTTCAAAAGCATGGAAGAGACAGTCTGCAACA 378  
319 TCAAAATATCGCTAGTGGCTGTACTTCAAAAGCATGGAAGAGACAGTCTGCAACA 378



QY 379 AGTATCGTGGTATTTGGGTAGCGCATTCGCTGGGATTTATGAAGAAGCAACGAACTTAGCA 438

Db 379 AGTACCGTGGTCTTGGGTAAACCCATTCGCTACATTTATGAAGGACACGGCATTTATTCG 438

QY 439 ATTGCTMAACAAGTGGGCTGAACCTCTAGCAAAACATTGCAAACTATAGAAACCGAAAGT 498

Db 439 TTTTCTMAACAAGTGGGCTGAACCGCTTCGCAACATTTGCTMACTGGAAGAAATCGAAGGT 498

QY 439 TGGATGTGCCCATTAAAGATCTGCCCCGCTCAGCAAAACATGTGTGATTCCTCGCTGTATGG 558

Db 439 TGGATGTGCTTCAAAATCTGCCCCGCTCGAAGAAACATGTGTGATTCCTGTGTGTATGG 558

QY 559 AAGCCCCAATCTGCACCGAAACGAAGTTATCTGGCAGGAAGAAAGGCTTGTGTGCGA 618

Db 559 AAGCCAAAACCTTTCACCAACGAGATTAATTGGAGCAGAGAAAGTACTTTCGAGTGCAG 618

QY 619 CGCTGCTCAGATTCTATCTGCTGCACAACTGTGTGACACGCGAGATGCTGCGAGTGT 678

Db 619 CTCTTATCCGGATTCGTTTGTGCTGACAGCCTGTGTGACGCGCATGGAGCTGCGAGCCT 678

QY 679 TCGCCAACCAAGTGTGCGCGCTCGCCATGAAGCCAAAAGAAATTGTGTGTCACCCCGAATA 738

Db 679 TCGCCAACCAAGTGTGTGTATGCACATGAGGCGAGAAAAATTCGTTAACATCTTTCAGCA 738

QY 739 AATATT 744

Db 739 AATATT 744

RESULT 3	CB036437	LOCUS	DEFINITION
	757 bp	mRNA	linear
	CB036437	TC.ad2.21605_TEXE1	Teladorsagia circumcincta adults library 2

ACCESSION CB0366437  
 VERSION CB0366437.1 GI:27755682  
 KEYWORDS EST.  
 SOURCE Teladorsagia circumcincta  
 ORGANISM Teladorsagia circumcincta  
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdida; Strongylida  
 Trichostrongyloidea; Haemonchidae; Oeserogastinae; Teladorsagia.  
 1 (bases 1 to 757)  
 AUTHORS Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Gulliano, D.,  
 Hall, N., Quayle, M. and Barrell, B.  
 TITLE Edinburgh University/Sanger Centre Nematode EST Project  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Blaxter ML

..... 3151

```
source
1..757
/organism="Teladorsagia circumcincta"
/mol_type="mRNA"
/db_xref="taxon:45464"
```

```

/clone="Tc_ad2_21G05"
/sex="mixed"
/dev stage="adult"
/clone.lib="Teladorsagia circumcincta adults library 2"
/note="Vector: pTribex2; Site 1: BamHI; Site 2: BamHI;
Teladorsagia circumcincta is a parasitic nematode. The
library was constructed from mRNA from Teladorsagia
circumcincta mixed adults."

```

Query Match	Score	DB	Length
Best Local Similarity	79.1%	Pred. No. 6,8e-133	
Matches	580	Conservative	0; Mismatches 147; Indels 6; Gaps 1

  

QY	19	GCGTGTGTAAGTCTGCTGTTCTCTCTGCGCCCTGTTCTCTATGCGAAGACGCTTTGTTG	78
DB	25	GCGAGTCAACATGGTGTCTTCTGCGCTTGTGCTCTTATGCGAAGACGCTTTGTTG	84
QY	79	TCCGAATGTGTAAGCCAAAGTGAACAGCGCGAGCGAGATTTTCCTCGATTTTCAACATGA	138
DB	85	CCCAAAAATCTTAAGCCAAACTGACAAAGCGAGGAGATTTTCTCGATTTTCAACATGA	144
QY	139	TGTTGTCGTAATATATAGCACTTGGAAATAGTGTGATTAATCTGSAACATTAATGACAGCGC	198
DB	145	TATTGTGTAACATAGCACTTGGAAAGAGCTGTGTAACCT-----TCAAGCATGATCC	198
QY	199	GATCATTTCTTGTCCTGACCTCGAACAATGTAACAAAGTGAACCTGGATTCGAACTTGGAGA	258
DB	199	GATGATCTTGTCCTGACCTCGAACAATGTAACAAAGTGAACCTGGATTCGAACTTGGAGA	258
QY	259	AGTGAAGCAACAACGATTTGGCGCATGTGAATGATCCCTACCGATTAATACAGCTGGC	318
DB	259	GAAAGCACTCAACGATTTGGCGCATGTGAATGATCCCTACCGATTAATACAGCTGGC	318
QY	319	TCAAATATCGCTAGATGCTGTACTTTCMAAGAAGTGAAGAGAAGACAGTTCGCAACA	378
DB	319	TCAAATATCGCAAGATGCTGTATTTATGCAACAGTGAAGAGAAGAGTTTTCAGACG	378
QY	379	AGTATCGTGTGATTTGGGTGAGCGCATCGCTGGATTTATGAAGGACAGAACTTGACCA	438
DB	379	AGTACCGGTGCTTGGGTGAACCCATTCCTTCAGTTCATGAAGGCGACGGGCACTTGATCG	438
QY	439	ATTGCTPAACAGATGGGGCTGAACCTCTTAGCAAACTGCAAACTATAGAAACCGAAAGGT	498
DB	439	TTTGCTPAACAGATGGGGCTGAACCGCTTGCAACATTTGCTAATCTGGAAGAAATTCGAAGGT	498
QY	499	TGATGTGCCCATTAAAGTCTGCCCCGCTGACGAAACAATGTAGTATCTCGTGTATG	558
DB	499	TGATGTGCTTAAAGAGTCTGCCCCGCTGAAAAAATGTGTGATCTCTGTGTATG	558
QY	559	AAGCCCCAAACTTGACCGAAGCAAGTATCTGGCAGGAAGAAAGGCTTGTGTGCGA	618
DB	559	AAGCCAAAACTTTTCAACCAACGAAGTATTTGGAGAGCAAGAAAGTATCTTGCAGATGCA	618
QY	619	CGCTGTCGCAATTCATCTGCTGCGAACAACCTGTGTGACACGCGAAGTGTGCGAGTGT	678
DB	619	CTCTTATCCGAATTCGTTTGTCTGTGACAGCTGTGTGACACGCTGTGAGAGCTTGGAGCT	678
QY	679	TCGCGACCAAGTGTGCGCTGCGCATGAAGCGAAAGAAATTTGTAGTCAACCCCGAATPA	738
DB	679	TCGCGACCAAGTGTGCTGATGACATGAGGCGAAGAAATTTCTGTATCAATCTTGACAG	738
QY	739	AATATGATGCAA 751	
DB	739	AATATGATTTAA 751	

<b>RESULT 4</b>			
CB038292			
<b>LOCUS</b>	760 bp	mRNA	linear
<b>DEFINITION</b>	TC ad2_4A12 TRXF1 Teladorsagia circumcincta adults library 2 Teladorsagia circumcincta cDNA clone TC ad_4A12 5' similar to CAD56659 Ancylostoma-secreted protein-like prote Obstreragia, mRNA		

sequence.  
CB038292 GI:27757537  
CB038292.1  
EST.  
Teladorsagia circumcincta  
Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongylidae; Haemonchidae; Oestertagiinae; Teladorsagia.  
1 (bases 1 to 760)  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guiliano, D.,  
Hall, N., Quayle, M. and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
Contact: Blaxter, M.  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRF1  
BACKWARD: T7PL  
Plate: 44 row: A column: 12  
Seq primer: TEXF1  
High quality sequence stop: 517.  
Location/Qualifiers  
1..760  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"  
/db\_xref="taxon:45464"  
/clone="TC\_ad2\_44A12"  
/sex="mixed"  
/dev\_stage="adult"  
/clone\_lib="Teladorsagia circumcincta adults library 2"  
/note="Vector: pTriplEx2; Site 1: BamHI; Site 2: BamHI;  
Teladorsagia circumcincta is a parasitic nematode. The  
library was constructed from mRNA from Teladorsagia  
circumcincta mixed adults."

ORIGIN

Query Match 62.1%; Score 474.2; DB 6; Length 760;  
Best Local Similarity 79.0%; Pred. No. 2.1e-132;  
Matches 579; Conservative 0; Mismatches 148; Indels 6; Gaps 1;

19 GGGCTGTAGTGTGTTCTCTGCGCCCTGTTCTCTTATGCGAAGAGGCTTTTGTG 78  
|||  
24 GGCAGTCATAGTGTGTTCTCTGCTTTGTCTCCATGCGAGGAGGCTTTTGTG 83  
|||  
79 TCCGAATAGTCTTAAAGCAAGTGAAGCGGAGGAGGATTTTCTTGATTTTCAATGA 138  
|||  
84 CCAAAAACCTTAAGCAAACTGACAAAGGAGGAGATTTTCTTGATTTTCAATGA 143  
|||  
139 TGTCTCCGAATATGACCTTGGAAATGTTGATTAAGTGAAGTAAATGACAGCG 198  
|||  
144 CATTCGTGGAACATGACCTTGGAAAGAGCTTGGTAAT-----TCAAGCTATTC 197  
|||  
199 GGTCAATCTTGTTCAGACTCAGAAAGTGAAGAGTGAAGTGAAGTGAAGTGAAG 258  
|||  
198 GGTGTTCTTGTTCAGACTCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 257  
|||  
259 AGTAGACAGCAACAGATTGCGCATGCAATGATCCCTTACCAATTAATCCAGCTGGC 318  
|||  
258 AAAAGAGCTCAACAGATTGCGCAATGACGTTCTCTCAATATGATCCAGCTGGC 317  
|||  
319 TCAAAATATGCTAGATGCTGTAATCTTCAAGACAGTGAAGAGAGAGCTTCAAGCA 378  
|||

Db 318 TCAAAATATGCAAGATGCTGTAATTAATGCAACAGTGAAGAGCAAGCTTTGAGACA 377  
Qy 379 AGTAGGGAATTTGGTGAAGCGCATGCTGGAATTAATGAAGGACGAACCTTGACA 438  
|||  
378 AGTACCGGTGCTTTGGTGAAGCGCATGCTGGAATTAATGAAGGACGAACCTTGACA 437  
|||  
439 ATTTGTCACCAAGTGGGCTGAACCTTGAAGCAATTCGAATTAATGAAGGACGAAGT 498  
|||  
438 TTTTGTACCAAGTGGGCTGAACCGTTGCAATTCGAATTAATGAAGGACGAAGT 497  
|||  
Qy 499 TGAATGTCCTTAATGATTCGCCCCGTCAGCAAAAGATGATTCCTGCGGTATGG 558  
|||  
Db 498 TGAATGTCCTTAATGATTCGCCCCGTCAGCAAAAGATGATTCCTGCGGTATGG 557  
|||  
Qy 559 AAGCCCAAACTTGCACGCAAGCAAGTATTCGCAAGAAAGGCTTGTGTGCGA 618  
|||  
Db 558 AAGCCCAAACTTGCACGCAAGCAAGTATTCGCAAGAAAGGCTTGTGTGCGA 617  
|||  
Qy 619 CGCTGTCAGATTCATTCGCTGCGACACACCTGTGACACGCGAGATGCTGCGAGTGT 678  
|||  
Db 618 CTCTTATCCGATTCGTTTGTGCTGTGACAGCTGTGTGACAGATGAGGCTGCGAGCT 677  
|||  
Qy 679 TCGCACCAAGTGTTCGCGCTGCGCATGAAGGCAAAAGATTTGATGACCCGAAATGA 738  
|||  
Db 678 TCGCACCAAGTGTTCGCGCTGCGCATGAAGGCAAAAGATTTGATGACCCGAAATGA 737  
|||  
Qy 739 AATATTCATGCA 751  
|||  
Db 738 AATATTCATGCA 750  
|||

RESULT 5  
LOCUS CB038112  
DEFINITION TC\_ad2\_42A03, TEXF1 Teladorsagia circumcincta adults library 2  
CAD56659 Ancylostoma-secreted protein-like prote Oestertagia, mRNA  
sequence.  
CB038112  
VERSION CB038112.1 GI:27757357  
KEYWORDS EST.  
SOURCE Teladorsagia circumcincta  
ORGANISM Teladorsagia circumcincta  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
Trichostrongylidae; Haemonchidae; Oestertagiinae; Teladorsagia.  
1 (bases 1 to 766)  
Blaxter, M.L., Parkinson, J., Whitton, C., Daub, J., Guiliano, D.,  
Hall, N., Quayle, M. and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project  
Unpublished (2000)  
Contact: Blaxter, M.  
Institute of Cell, Animal and Population Biology  
University of Edinburgh  
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9  
3JT, UK.  
Tel: +44 131 650 6760  
Fax: +44 131 670 5450  
Email: mark.blaxter@ed.ac.uk  
The library was prepared by Dr Diane Redmond and Dr David Knox,  
Moredun Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.  
Sequencing was performed by the Pathogen Sequencing Unit, Sanger  
Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell).  
PCR Primers  
FORWARD: TEXPCRF1  
BACKWARD: T7PL  
Plate: 42 row: A column: 03  
Seq primer: TEXF1  
High quality sequence stop: 516.  
Location/Qualifiers  
1..766  
/organism="Teladorsagia circumcincta"  
/mol\_type="mRNA"

FEATURES  
source



Db	319	TCAAAATRTGCGAAGATGGCTGTATTTATGCCACAGTGAAGAAAGCAAGTTTGAGACA	378
Qy	379	AGTATCGTGATATTGGGTGAGCGCATCGCTGGGATTATGAAAGGACAGAACTTGAACCA	438
Db	379	AGTACCGGTGCTTGGGTAAACCCATCCGTCAGATTATGAAAGGCAAGGACCTTGATCG	438
Qy	439	ATTGTCTAACAGTGGGCTGAAACCTCTTGAAGAAACATTCGAAACCTATAGAAACCGAAGGT	498
Db	439	TTTGTCTAACAGTGGGCTGGAACGGCTTGCAACCTCTTAACTTGAAAAATTCGAAAGT	498
Qy	499	TGATGTGCCATTAAGATCTGCCCCGCTCAGCAAAACATGTAGTATCTGCGTGTATGG	558
Db	499	TGATGTGCTTAAACAAATCTGCCCCGCTGAAAAACATGTGCTATCTGTGTGTATGG	558
Qy	559	AAGCCCCAACTTGACCGAAAGAAATTATCTGGCAGGAAGAAAGGCTTGTGTGCGA	618
Db	559	AAGCAAAABACTTTCACCAAAGAAAGTATTTGGGACGAGGAAGTACTTGGAGATGCAA	618
Qy	619	CGCTGCTCAAGTATCTGTCTGCTGGACAACTGTGTGACACGCGAGATGCTGCCAAGTGT	678
Db	619	CTCTTATCCGGATTCGTTTGTCTGTGACAGCTCTGTGTGACAGCATGAGAGCTGCAGGCT	678
Qy	679	TCGCCACAGTTCGCGGTGCGCANTGAAGCGAAAGAAATTGTAGTCAACCCCGAATTA	738
Db	679	TCGCCACAGTGTGTGTATGACAGTAGGCGAGAAAAATTTCTGTTAACATCTTGAGCAA	738
Qy	739	AATATT 744	
Db	739	AATATT 744	

RESULT 7					
CB036975	CB036975	755 bp	mRNA	linear	BST 15-JAN-2003
LOCUS					
DEFINITION	Tc.ad2.28D02_TEXE1 Teladorsagia circumcincta adultes				library 2

sequence.

ACCESSION	CB036975
VERSION	CB036975.1
	GI:27756220

KEYWORDS	EST.
SOURCE	Teladorsagia circumcincta
ORIGIN	Belgium

ORGANISM  
*Teleduobagia circumcincta*  
Eukaryota; Metazoa; Chromadorea; Rhabdida; Strongylida  
Trypanostomatoidae. Haemonchidae. Ostertagiinae: *Teleduobagia*.

REFERENCE  
1 (bases 1 to 755)  
Blayter M.L., Parkinson, J., Whittton, C., Daub, J., Guiliano, D.,

**TITLE** Hall, N., Quayle, M. and Barrell, B.  
Edinburgh University/Sanger Centre Nematode EST Project

Journal  
Unpublished (2000)  
Contact: Blaxter ML

Institute of Cell, Animal and Population Biology  
University of Edinburgh

Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK.

Tel: +44 131 650 6760  
Fax: +44 131 670 5450

Email: [mark.blaxter@ed.ac.uk](mailto:mark.blaxter@ed.ac.uk)  
The library was prepared by Dr Diane Redmond and Dr David Knox

Moredon Research Institute, Midlothian, UK.  
PCR was performed by Ye Jieru, ICAPB, University of Edinburgh.

Sequencing was performed by the Patogen Sequencing Unit, Safer Centre, Cambridge, UK (Neil Hall, Mike Quail & Bart Barrell)

PCR Primers  
FORWARD: TEXPREF1  
REVERSE: TREF1

BALMARD: 1/PL  
Plate: 28 row: D column: 02  
Sec primer: TTYE1

seq primer: left  
High quality sequence stop: 516.  
Location/Qualifier

```

1. .755
/organism="Teladorsagia circumcincta"

```

```

/mol_type="rRNA"
/db_xref="taxon:45464"
/clone="TC ad2_28D02"
/sex="mixed"
/dev_stage="adult"
/clone_lib="Teladorsagia circumcincta adults library 2"
/note="Vector: pTriplEx2; Site_1: BamHI; Site_2: BamHI; Teladorsagia circumcincta is a parasitic nematode. The library was constructed from mRNA from Teladorsagia circumcincta mixed adults."

```

Query Match	62.0%;	Score 473.2;	DB 6;	length 755;
Best Local Similarity	79.3%;	Pred. No. 4.2e-132;		
Matches 575;	Conservative 0;	Mismatches 144;	Indels 6;	Gaps 1.

QY 20 CCTGTTGAGTGTGCTGTTCTCTGGCCCTGTTCTCTTATCCGAAGCAGGCTTTGTGT 79

Db 27 GCTGTCAAGTTTGTCTCTCTCTGCTCTGTCTCTTAAGCCGAAGCAGGCTTTGTGTG 86

QY 80 CCGATAGTCTTAAGCCCAAGTAGACACGCGCGAGCATTTTCTCGATTTTCAATGAT 139

Db 87 CCAAAAATCTGGGCCAAACGACAAACGCAAGCATTTTCTCGATTTTCAATGAT 146

QY 140 GTTCGTCGAATATATAGCACTTGGAATGTTGATTAACCTGACAGTAAATGACAGCG 199

Db 147 ATTGCTCGGAACATAGCACTTGGAATGAGCTTGTAACCT-----TCACGACGAGATCG 200

200 GTCATTTCTGGTCAGCTCAGAACATGTACAAGTGCCTGGATTGCACTTGGAAGA 259

Db 201 GTATTCTTGGTCAAGCTCAAAACATGTACAAGCTGATTGGACTGTGACTGGACAA 260

260 GTAGCAGCAACAGATTGGCCATGATCCCTACCGATTAATATCAGCCTGGCT 319

Db 261 AAAGCAGCCAGCAGATTGGCGAATGCACGGTTCCTTAACCAATAGATCCAGCCTTGCT 320

320 CAAATATCGCTAGTGGCTGCTACTTCAAAGA CAGTGAAGAGACAGTTCTGCAACA 379

Db 321 CAAATATCGCAGATGGCTGTATTATGCCACAGTGAAGAACAAGTTTGAGCAA 380

380 GATCGTGATTTGGGTGACGCATCGCTGGATTATGAAAGCAGCAACTTGACCA 439

Db 381 GTACCGTCTTGGGTAAACCCACACGTACATTTAAGAGGGACGGCATTGATCG 440

[illegible][illegible]

D500 GGAAGTGTCTGTTGGATGACGCGCCTGAAAGAACATGCATCCCATCCCTGGTGATGGG A5600

560 AACGCCCCAAGCTTGGACCCGGAACGGAAGTATCTTGGCAGGAGAGGAAGGCTTGTGTGTGCGAC 615

561 AGCTCAAAACCTTTGACCAACGAAAGTACTGGAGCAGGATACTTGGAGTGCAC 620

620 GCTGCTCAGATTCAATCTGCTGCCGACACCTGTGTGACACGCCAGATGCTCCGAGTCTT 679

621 TCTTATCCGAGTTCGTTTGTCTGTGACAGCCCTGTGTGACAGCAGTCGTGCGAGCCTT 680

680 CGCCACCACTTGTGCGCGTCCGCATGAAGGAAAGAATTGTGTA CTCACCCGAATAA 739

681 CGCCACCAAGTGTGTTGATCGACATGAGCGGAGAAAAATTCTGGTAAACATCTTGACAAA 740

QY 740 ATATT 744

Db 741 ATATT 745

RESULT 8					
CB038323					
LOCUS	751 bp	mRNA	linear	EST 15-JAN-2007	
DEFINITION	Tc_ad2_44D11_TEXF1 Teladorsagia circumcincta adults		library 2		

```

seq primer: TEXP1
High quality sequence stop: 517.
location/Qualifiers
source            1       754
FEATURES

```















GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2006, 19:02:37 ; Search time 229 Seconds  
(without alignments)  
711.690 Million cell updates/sec

Title: US-10-527-771-10

Perfect score: 1253  
Sequence: 1 MSAVVAVVAVLALFSAEAG.....DNLCDTRDAASVHQCCASP 231

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80:\*  
1: uniprot\_gprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1253	100.0	231	2	Q81FT6 OSTOS
2	638.5	51.0	236	2	Q81FT6 OSTOS
3	277	22.1	424	1	ASP ANCA
4	274	21.9	425	2	Q81FX0 9B1LA
5	274	21.9	425	2	Q81FX0 9B1LA
6	263	21.0	424	2	Q81FX0 9B1LA
7	260	20.8	424	2	Q81FX0 9B1LA
8	211	16.8	424	2	Q81FX0 9B1LA
9	211	16.8	424	2	Q81FX0 9B1LA
10	205	16.4	425	2	Q81FX0 9B1LA
11	195.5	15.6	424	2	Q81FX0 9B1LA
12	191	15.2	424	2	Q81FX0 9B1LA
13	186.5	14.9	424	2	Q81FX0 9B1LA
14	178.5	14.2	424	2	Q81FX0 9B1LA
15	176	14.0	424	2	Q81FX0 9B1LA
16	167	13.3	424	2	Q81FX0 9B1LA
17	166	13.2	424	2	Q81FX0 9B1LA
18	164	13.1	424	2	Q81FX0 9B1LA
19	161.5	12.9	424	2	Q81FX0 9B1LA
20	160	12.8	424	2	Q81FX0 9B1LA
21	158.5	12.6	424	2	Q81FX0 9B1LA
22	156	12.5	424	2	Q81FX0 9B1LA
23	152.5	12.2	424	2	Q81FX0 9B1LA
24	151	12.1	424	2	Q81FX0 9B1LA
25	150	12.0	424	2	Q81FX0 9B1LA
26	146	11.7	424	2	Q81FX0 9B1LA
27	145	11.6	424	2	Q81FX0 9B1LA
28	142.5	11.4	424	2	Q81FX0 9B1LA
29	142.5	11.4	424	2	Q81FX0 9B1LA
30	138	11.0	424	2	Q81FX0 9B1LA
31	135.5	10.8	424	2	Q81FX0 9B1LA

32	135	10.8	246	2	Q9N313 CAEBL
33	135	10.8	459	2	Q45132 HAECO
34	132.5	10.6	208	2	Q9N5N4 CAENORHABDI
35	132.5	10.6	212	2	Q61FT8 CAEBR
36	131.5	10.5	224	2	Q86FX1 ONCHOE
37	130.5	10.4	251	2	Q8FVN7 RATV
38	130	10.4	209	2	Q62317 CAEBR
39	129	10.3	206	2	Q62317 CAEBR
40	128.5	10.3	220	2	Q16854 ONCHOE
41	127.5	10.2	361	2	Q624A6 CAEBR
42	125.5	10.0	213	2	Q60KMS CAENORHABDI
43	124.5	9.9	224	2	Q9GPN4 ONCHOE
44	124.5	9.9	303	2	Q724U9 HUMAN
45	124.5	9.9	446	2	Q6DCB3 HUMAN

## ALIGNMENTS

RESULT 1	ID	Q81FT6 OSTOS	PRELIMINARY	PRT	231 AA
AC	Q81FT6	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)				
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)				
DE	Angiostoma-secreted protein-like protein.				
GN	Name=asap2;				
OS	Ostertagia ostertagi.				
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;				
OC	Trichostrongyloidea; Haemonchidae; Ostertagiinae; Ostertagia.				
OX	NCBI_TaxID=6317;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RX	MEDLINE=22593580; PubMed=12706806; DOI=10.1016/S0166-6851(03)00044-6;				
RA	Geldhof P.B., Vercauteren I., Gevaert K., Staes A., Knox D.,				
RA	Vandekerckhove J., Vercauteren J., Claeys R.,				
RT	"Activation-associated secreted proteins are the most abundant				
RT	antigens in a host protective fraction from Ostertagia ostertagi."				
RL	Mol. Biochem. Parasitol. 128:111-114(2003).				
DR	EMBL; AJ515523; CAD56659.1; -; mRNA.				
DR	GO; GO:0005576; C:extracellular region; IEA.				
DR	InterPro; IPR001283; Allrgn_VS/Tp1.				
DR	Pfam; PF00188; SCP; 1.				
DR	SMART; SM00198; SCP; 1.				
SQ	SEQUENCE 231 AA; 25504 MW; 881043E42AF07081 CRC64;				
Query Match	100.0%; Score 1253; DB 2; Length 231;				
Best Local Similarity	100.0%; Pred. No. 5.3e-105;				
Matches	231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MSAAVVAVVAVLALFSAEAGFCPCNSLSQSDSAQIFLDFNDVNRNALANGLINVTN	60		
DB	1	MSAAVVAVVAVLALFSAEAGFCPCNSLSQSDSAQIFLDFNDVNRNALANGLINVTN	60		
QY	61	ADAVILGPAQMYVVDMDNLEVAADQIARCNPLPINTSLAQTIAWLPKSEESTV	120		
DB	61	ADAVILGPAQMYVVDMDNLEVAADQIARCNPLPINTSLAQTIAWLPKSEESTV	120		
QY	121	LQGVSWVWVVASLGFMTKTKLDQFANQWAEPLANIYANRNRKVCAGHKICPAQNMVVC	180		
DB	121	LQGVSWVWVVASLGFMTKTKLDQFANQWAEPLANIYANRNRKVCAGHKICPAQNMVVC	180		
QY	181	VYGSBKAPNEVINOEGKACVCDARPSFCCDNLCDFRDAASVHQCCASP 231			
DB	181	VYGSBKAPNEVINOEGKACVCDARPSFCCDNLCDFRDAASVHQCCASP 231			
RESULT 2	Q81FT6 OSTOS				
AC	Q81FT6	01-JUN-2003 (TREMBLrel. 21, Created)			
DT	01-JUN-2003 (TREMBLrel. 21, Created)				

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Ancylostoma-secreted protein-like protein (fragment).  
 GN Name=OSP-2;  
 OS Oostertagia ostertagi.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 CC Trichostrongylidae; Haemonchidae; Oostertaginae; Oostertagia.  
 NC NCB1\_TaxID=6317;  
 RX NCBI\_TaxID=6317;  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=22503448; PubMed=12615319; DOI=10.1016/S0166-6851(02)00274-8;  
 RA Vercauteren I.J.R., Geldhof P., Pelelers I., Claerebout E., Bex G.,  
 RA Vercauteren J.;  
 RT "Identification of excretory-secretory products of larval and adult  
 RT Oostertagia ostertagi by immunoscreening of cDNA libraries.";  
 RL Mol. Biochem. Parasitol. 126:201-208(2003).  
 DR EMBL; AJ130812; CAD23183.1; -; mRNA.  
 DR GO; GO:0005576; C:extracellular region; IEA.  
 DR InterPro; IPR001283; Allrgn\_V5/Tpx1.  
 DR Pfam; PF00188; SCP; 1.  
 DR SMART; SM00198; SCP; 1.  
 FT NON TER 1  
 SQ SEQUENCE 236 AA; 26228 MW; DB52781B8353A698 CRC64;  
  
 Query Match 51.0%; Score 638.5; DB 2; Length 236;  
 Best Local Similarity 52.5%; Pred. No. 2e-49; Indels 11; Gaps 4;  
 Matches 125; Conservative 36; Mismatches 66;  
  
 QY 1 MSAVVAVVAV-LTALFSYVAGFCCPNSISQSDSARQIFLDPHNDVRRNIATLNGILINWT 58  
 DB 1 MOTALGIALVAVLTSTNTAGFCPPADLNQDEBARKITLDPHNVRHDIAGASPLNLT 60  
 QY 59 VNADAV---ILGPAQNMKYKVDMDCNLEBVAQAQIAPCNDPLPINTSLAQNIAKMLYFED 114  
 DB 61 ---GAVQRMNVIGPAAKMYRMMDNDCLBAKAKAMTWPTTPIPIPTSIQNTAOWILTFQ 117  
 QY 115 SEEEVTVOQVSWYVWSASLGPMKGTGLDPPANQMAEPLANIYRRKRGCAHKTC--PA 172  
 DB 118 SOEXEVLITQTPMSWVYASLRNLQPTTEANIYMQIRPLSINAMWOLKVGCAHKCKEPT 177  
 QY 173 QQNMVVCYVSGPKLAPNEVIMQEGKACVCDARPDSCDNLCDTFDAASVRRQCCAS 230  
 DB 178 GTNMVVSCHVGGSEVLQDNEVMDKPTCCNMAVPSFCNNICDTIAATLQNPCKS 235  
  
 RESULT 3  
 ASP\_ANCCA STANDARD; PRT; 424 AA.  
 ID ASP\_ANCCA  
 AC Q1637;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Ancylostoma secreted protein precursor.  
 GN Name=ASP;  
 OS Ancylostoma caninum (Dog hookworm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 CC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
 NC NCB1\_TaxID=29170;  
 RX [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Haddon J.M., Jones B.F., Hoffman D.R., Hotez P.U.;  
 RA MEDLINE=96215086; PubMed=8636085; DOI=10.1074/jbc.271.12.6672;  
 RT "Cloning and characterization of Ancylostoma-secreted protein. A novel  
 RT protein associated with the transition to parasitism by infective  
 RT hookworm larvae.";  
 RL J. Biol. Chem. 271:6672-6678(1996).  
 CC -1- FUNCTION: Associated with the transition to parasitism by  
 CC infective hookworm larvae.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).  
 CC -1- SIMILARITY: Belongs to the CRISP family.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 DR EMBL; U26187; AAC47001.1; -; mRNA.  
 DR HSSP; P04284; ICBE.  
 DR InterPro; IPR001283; Allrgn\_V5/Tpx1.  
 DR InterPro; IPR002413; V5\_allergen.  
 DR Pfam; PF00188; SCP; 2.  
 DR PRINTS; PR00838; V5ALLERGEN.  
 DR PRINTS; PR00837; V5TPXLIK.  
 DR PRODOM; PD000542; Allrgn\_V5/Tpx1; 1.  
 DR SMART; SM00198; SCP; 2.  
 DR PROSITE; PS01009; CRISP\_1; 1.  
 DR PROSITE; PS01010; CRISP\_2; FALSE\_NEG.  
 KW Signal.  
 FT SIGNAL 1  
 FT CHAIN 19 424 Potential.  
 FT CHAIN 19 424 Ancylostoma secreted protein.  
 SQ SEQUENCE 424 AA; 45735 MW; 4BC8295F5D3035F9 CRC64;  
  
 Query Match 22.1%; Score 277; DB 1; Length 424;  
 Best Local Similarity 30.9%; Pred. No. 1.9e-16;  
 Matches 76; Conservative 35; Mismatches 97; Indels 38; Gaps 11;  
  
 QY 1 MSAVVAVVAVLAL-----FSYVAGFCCPNSISQSDSARQIFLDPHNDVRRNIATLNGILIN 56  
 DB 1 MFSPIYVITITAFCDASPARDFGCGNS-GITDKORQATLDPHNNARRVAKG----- 54  
 QY 57 WTNADAVITLPPAQMRYKVDMDCNLEBVAQAQIAPCNDPLPINTSLAQNIAKMLY---FK 113  
 DB 55 -VEDSNGKLNAPKMYKTLSDCAMEQQLQDAIOSCPBAPAGIQCAQNVMSWSSSGGP 113  
 QY 114 DSEERTVOQVSWYVWSASLGPMKGTGLDQ-----FANQWABPLANIYRRKRGVCA 166  
 DB 114 DSVYV-IEQTLSGWSGK---KNGVGPDKRYNGSGLPV-----FSMNVY3ETTKLGCA 163  
 QY 167 HKICPAQNMVVSCHVSGPKLAPNEVIMQEGKACV---CDARPDSCDNLCDT-RDAA 221  
 DB 164 YKVGCTK--LAVSCITNGVGTITNPMMETGQACTGADCSYKNGSGEDGLCTGPDVP 221  
 QY 222 SVRQOC 227  
 DB 222 ETNQOC 227  
  
 RESULT 4  
 O816X0\_9BILA PRELIMINARY; PRT; 425 AA.  
 ID O816X0\_9BILA  
 AC O816X0;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Secreted-protein 1.  
 GN Name=aspl1;  
 OS Ancylostoma ceylanicum.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 CC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
 NC NCB1\_TaxID=53326;  
 RX [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Zhan B., Liu Y., Hotez P.J., Haddon J.;  
 RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AT135548; AN114402.1; -; mRNA.  
 DR GO; GO:0005576; C:extracellular region; IEA.  
 DR InterPro; IPR001283; Allrgn\_V5/Tpx1.  
 DR InterPro; IPR002413; V5\_allergen.  
 DR Pfam; PF00188; SCP; 2.  
 DR PRINTS; PR00838; V5ALLERGEN.  
 DR PRINTS; PR00837; V5TPXLIK.  
 DR PRODOM; PD000542; Allrgn\_V5/Tpx1; 1.  
 DR SMART; SM00198; SCP; 2.  
 DR PROSITE; PS01009; CRISP\_1; 1.  
 DR PROSITE; PS01009; CRISP\_2; 1.  
 SQ SEQUENCE 425 AA; 45879 MW; B995D5B16B0C2250 CRC64;

Query Match 21.9%; Score 274; DB 2; Length 425;  
 Best Local Similarity 30.5%; Pred. No. 3.5e-16;  
 Matches 74; Conservative 34; Mismatches 101; Indels 34; Gaps 9;

2 SAAYVAVLAL-----FSYAEAGFCCPNLSGSDSARQIFLPHNDVRNIALG-----N 52  
 Db 3 SSVVAVISVLTVAFCDDASPVKASFGCSNS-GITDSDRQAFLEPHNNARRRVAQGVEDNKS 61  
 Qy 53 GLINTVNAVAVILGPAQNMVYVMDQNLBEVAAQOIAPCNDPLPINTSLAQNIRMLY- 110  
 Db 62 GRLN-----PAKMYTLFMDCKMEQQLQDAIGSCPGGAGIGQVAGNIIISWSGS 110  
 Qy 111 -YFKDSEETVLAQVSWYVWSASLGFPMKGTQLDQFANQWAEPLANIANRNRKVGCAHKI 169  
 Db 111 GGFPPNSEKINTSLASWAGAKNGVASDNK---YTGGLYAFSNMVPSETTKLGCAHYV 167  
 Qy 170 CPAQNMVSCVYSGPKLAPNEVIMQEGKACV---CDARPDSFCCDNLCDT-RDAASV 224  
 Db 168 CGTK--LTLSCITNGIGYMTGA PMMETGQACKAGADCTTFKNSGCEGLCTKGADVPETN 225  
 Qy 225 HQC 227  
 Db 226 QQC 228

## RESULT 5

077153\_9B1LA PRELIMINARY; PRT; 425 AA.  
 AC 077153;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ancylostoma-secreted protein 1 precursor.  
 GN Name=aspl;  
 OS Ancylostoma duodenale.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
 CX NCBI\_TaxID=51022;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Shanghai;  
 RX MEDLINE=99151774;  
 RA Bin Z., Hawdon J., Qiang S., Hainan R., Huigang Q., Wei H.,  
 RA Shu-hua X., Tiehua L., Xing G., Zheng F., Hotez P.;  
 RT "Ancylostoma secreted protein 1 (ASP-1) homologues in human  
 hookworms";  
 RL Mol. Biochem. Parasitol. 98:143-149 (1999).  
 DR EMBL; AF077402; AAD1339.1; -, mRNA.  
 DR GO; GO:0005576; C:extracellular region; IEA.  
 DR InterPro; IPR001283; Allrgn V5/Tp1.  
 DR InterPro; IPR002413; V5\_allergen.  
 DR Pfam; PF00188; SCP\_2; V5\_allergen.  
 DR PRINTS; PR00838; V5ALLERGEN.  
 DR PRINTS; PR00837; V5TPXLKE.  
 DR SMART; SM00198; SCP\_2.  
 DR PROSITE; PS01009; CRISP\_1; 1.  
 KW SIGNAL.  
 FT CHAIN 1 19 Potential.  
 FT SIGNAL 20 425 ancylostoma-secreted protein 1.  
 SQ SEQUENCE 425 AA; 45821 MW; 1F6B9D7E62EBDAB CRC64;  
 Query Match 21.9%; Score 274; DB 2; Length 425;  
 Best Local Similarity 30.3%; Pred. No. 3.5e-16;

Matches 74; Conservative 37; Mismatches 97; Indels 36; Gaps 11;

2 SAAYVAVLAL-----FSYAEAGFCCPNLSGSDSARQIFLPHNDVRNIALG-----N 52  
 Db 3 SSVVAVISVLTVAFCDDASPVKASFGCSNS-GITDSDRQAFLEPHNNARRRVAQGVEDNKS 61  
 Qy 53 GLINTVNAVAVILGPAQNMVYVMDQNLBEVAAQOIAPCNDPLPINTSLAQNIRMLY- 111  
 Db 62 GRLN-----PAKMYTLFMDCKMEQQLQDAIGSCPGGAGIGQVAGNIIISWSGS 110  
 Qy 112 -FKDSEETVLAQVSWYVWSASLGFPMKGTQLDQFANQWAEPLANIANRNRKVGCAHKI 168  
 Db 111 GGFPPNSEKINTSLASWAGAKNGVASDNK---YTGGLYAFSNMVPSETTKLGCAHYV 166  
 Qy 169 ICPAQNMVSCVYSGPKLAPNEVIMQEGKACV---CDARPDSFCCDNLCDT-RDAASV 223  
 Db 167 VCGTK--MATSCITNGIGYMTGA PMMETGQACKAGADCTTFKNSGCEGLCTKGADVPET 224  
 Qy 224 RHQC 227  
 Db 225 NQC 228

## RESULT 6

076744\_NECAM PRELIMINARY; PRT; 424 AA.

AC 076744;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ancylostoma secreted protein 1 precursor.  
 GN Name=aspl;  
 OS Necator americanus (Human hookworm).  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
 OC Ancylostomatidae; Ancylostomatidae; Bunostominae; Necator.  
 CX NCBI\_TaxID=51031;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Shanghai;  
 RX MEDLINE=96215086; PubMed=8636085; DOI=10.1074/jbc.271.12.6672;  
 RA Hawdon J.M., Jones B.F., Hoffman D.R., Hotez P.J.;  
 RT "Cloning and characterization of Ancylostoma-secreted protein. A novel  
 protein associated with the transition to parasitism by infective  
 hookworm larvae";  
 RL J. Biol. Chem. 271:6672-6678 (1996).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=Shanghai;  
 RX MEDLINE=99151774;  
 RA Bin Z., Hawdon J., Qiang S., Hainan R., Huigang Q., Wei H.,  
 RA Shu-hua X., Tiehua L., Xing G., Zheng F., Hotez P.;  
 RT "Ancylostoma secreted protein 1 (ASP-1) homologues in human  
 hookworms";  
 RL Mol. Biochem. Parasitol. 98:143-149 (1999).  
 DR EMBL; AF079521; AAD13340.1; -, mRNA.  
 DR HSP; P04284; 1CRF.  
 DR GO; GO:0005576; C:extracellular region; IEA.  
 DR InterPro; IPR001283; Allrgn V5/Tp1.  
 DR InterPro; IPR002413; V5\_allergen.  
 DR Pfam; PF00188; SCP\_2; V5\_allergen.  
 DR PRINTS; PR00838; V5ALLERGEN.  
 DR PRINTS; PR00837; V5TPXLKE.  
 DR SMART; SM00198; SCP\_2.  
 DR PROSITE; PS01009; CRISP\_1; 1.  
 KW SIGNAL.  
 FT CHAIN 1 18 Potential.  
 FT SIGNAL 19 424 ancylostoma secreted protein 1.  
 SQ SEQUENCE 424 AA; 45742 MW; B7B2F95F9BA9F CRC64;  
 Query Match 21.0%; Score 263; DB 2; Length 424;  
 Best Local Similarity 30.5%; Pred. No. 3.4e-15;  
 Matches 73; Conservative 36; Mismatches 106; Indels 24; Gaps 9;





Best Local Similarity 31.0%; Pred. No. 2.1e-09;  
Matches 54; Conservative 34; Mismatches 57; Indels 29; Gaps 8;

```

Qy 22 CCPSNLSQSDSARQJFLDPSNDVRRJIALG-----NGLINWTVNADAVILGPKNONYKXD 76
Db 60 CNESTITLOQ-QEILITTHNELRRSLAIGKOKRKEGLNM-----SARNYTKLD 106
Qy 77 WDCNLEBVAQOIAPCND---PLPINTSLAONIARMLYFQDSHEEYV-LQOVSWMYVSAS 132
Db 107 WDCELAALAAWNASASCPQHFMPOSIVGSNAQLFKRYTYTTFDGHDSIVHHRNANKYMWQ- 165
Qy 133 LGFMKGTJLDQ-----FANQMAEPLIANIAYNRKVCJCAHKICPAQONMVVSCYV 182
Db 166 -GEERBNE-DQQRKRFPRNRNYFGMANNAKGYTRVGCSTYMCDDTSALFTCIY 217

```

RESULT 12  
086GK5 AN

```

086GK5 ANCCA
ID 086GK5 ANCCA PRELIMINARY; PRT; 424 AA.
AC 086GK5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Secreted protein 5.
GN Name=asps;
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatoida; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22788807; PubMed=12906874; DOI=10.1016/S0020-7519(03)00111-5;
RA Zhao B., Liu Y., Badamchian M., Williamson A., Peng J., Loukas A.,
RA Hawdon J.M., Hotez P.J.;
RT "Molecular characterization of the Ancylostoma-secreted protein family
RL from the adult stage of Ancylostoma caninum.";
RL Int. J. Parasitol. 33:897-907(2003).
DR EMBL; AY217006; AA063577.1; -; mRNA.
DR HSSP; P04284; ICPE.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; A1lrgn_V5/Tpx1.
DR Pfam; PF00188; SCP_2.
DR PRINTS; PR00837; V5TPXLIKE.
DR ProDom; PD000542; A1lrgn_V5/Tpx1; 1.
DR SMART; SM00198; SCP_2.
SQ SEQUENCE 424 AA; 46227 MW; E92AF9059D61F28D CRC64;

```

Query Match	15.28;	Score 191;	DB 2;	Length 424;
-------------	--------	------------	-------	-------------

Best Local Similarity 29.5%; Pred. No. 1.1e-08;  
Matches 62; Conservative 31; Mismatches 83; Indels 34; Gaps 11

QY 23 CENSLQSQSSAQCIFLDFHNDVPRNRLALNGNLNWTVMNDVAITLSPQAQNNKYKDMQCNLE 82  
 Db 21 CEGN-DLTLAERTLLRVNHSIRRELA--QGVAN---NTHGKLPAGKNTYRNRVCELE 74  
 QY 83 EYA--AQIAPCNDPLPINTSLAQNTARMLY-----FKDSEETVLQGVSWYMWASAL 133  
 Db 75 QALIDASQTF-CGASLEBPQKGCQNTQAYTTSPITARPKNDDLBDNAVKQ---WYLPITY 130  
 QY 134 GFMRKTKLDQFANQMAEP-----LANTANYNRIRYGCANHKICPAQAQNNVVSCTYGSPEGLAP 189  
 Db 131 GQRDA-----ANKFTDPRLYTFANLADYDKNTALMGCHYACQGPDRIVISCMYNN--VVPD 183  
 QY 190 NEYIMQEGKACV-----CDARPDSPCCDNL 215  
 Db 184 NNAVIEPGTACVDADCTTYPQSTCKDSL 213

RESULT 13	
Q86GK4_ANCCA	
ID Q86GK4_ANCCA	PRELIMINARY;
AC Q86GK4;	PRT;
DT 01-JUN-2003	451 AA
(TREMBLrel. 24, Created)	

DT 01-JUN-2003 (TrEMBLrel. 24, last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
DS Secreted protein 6.

GN Name=asp6;  
OS Ancylostoma caninum (Dog hookworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.  
OX NCBI\_TaxID=29170;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=22788807; Pubmed=12906874; DOI=10.1016/S0020-7519(03)00111-5;  
RA Zhan B., Liu Y., Badamchian M., Williamson A., Feng J., Loukas A.,  
RA Haddon J.M., Hotez P.J.;  
RT "Molecular characterization of the Ancylostoma-secreted protein family  
RT from the adult stage of Ancylostoma caninum.";  
RL Int. J. Parasitol. 33:897-907(2003).  
DR EMBL; AY217007; AA063578.1; '-' mRNA.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR InterPro; IPR001283; Align\_V5/Trpt.  
DR Pfam; PF00188; SCP\_2  
DR SMART; SM00198; SCP\_2  
SQ SEQUENCE 451 AA; 48760 MW; 9704247C59C294FB CRC64;

Query Match	14.9%	Score 186.5;	DB 2;	Length 451;
Best Local Similarity	27.1%;	Pred. No. 3e-08;		
Matches	59;	Conservative	39;	Mismatches 95;
				Indels 25;
				Gaps 9

```

Oy 7 VAVILALFSYEA--GCCPNSLSOSDSARQFLDPHNDVRRNIALANGLIMTVNADAVI 65
Db 5 ILVLVAILGIAHADFQCM--FKSTDLREHLKSLINLRKKIADG-----SAENKSGK 57
Oy 66 LGPAONMYKVDMDCNLEEVAAQIAPNCDDPLPINTSLAQNIAWLYFYDSEET--VLQQ 123
Db 58 CPQGNRIKTLMDDELEIKQAQAVDQCKPNVEBAGYSQILKK--VXSTCDPTKVLKKQ 114
Oy 124 VSWTVVASLGFMKCTKLDQFAN--QMAEPLANTINYRRKYGCAHKICPAQONNVSGV 181
Db 115 IEAWMTKS----VKDAGVDNPPNNKQGLJEDPAKLANGATKIGCAQKNC--NEQLYVACV 168
Oy 142 YGSPKLAPNEVIMQEGKAC---VCDARPDSPCCDNLK 215
Db 169 INEPAPVAGMPITYEVAGAGCNSKDDCTTYLQSKCSNKCVC 206

RESULT 14
P90959_CABEL.
ID P90959_CABEL PRELIMINARY; PRT: 221 AA.
AC P90959;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, last sequence update)
DI 01-JUN-2003 (TREMBLrel. 24, last annotation update)
DE Hypothetical protein T05A10.4.
GN ORName=T05A10.4;
OS Caenorhabditis elegans.
OC Rukariyota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99063613; PubMed=9851916;
RT The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; Z68108; CAA32137.1; -; Genomic_DNA.
DR PIR; T24494; T24494.
DR Ensemble; T05A10.4; Caenorhabditis elegans.
DR WormBase; WBGene00011461; T05A10.4.
DR WormPep; T05A10.4; CE13187.
DR GO; GO:0005576; C:extracellular region; IEA.
DR InterPro; IPR001283; Altrgm_V5/Ipx1.
DR InterPro; IPR002413; V5_allergen.

```



DR Pfam, PF00188, SCP, 1.  
DR PRINTS; PR00838; VSALLERGEN.  
DR SMART; SM00198; SCP, 1.  
KW Complete proteome: Hypothetical protein.  
SQ SEQUENCE 221 AA; 25494 MW; FE2EC5284DB06f89 CRC64;

Query Match 14.2%; Score 178.5; DB 2; Length 221;  
Best Local Similarity 29.3%; Pred. No. 7.3e-08;  
Matches 51; Conservative 35; Mismatches 59; Indels 29; Gaps 8;

QY 22 CPNLSQSQDSARQIFLDNDVRRNIALG-----NGLINWTVNADAVILGPAQNMVYKVD 76  
DB 61 CKKSTITDQ--QEILITTHNELRSILAGCKRNKRLMN-----GARNMYKLD 107  
QY 77 WDCNLEEVAAQQLAPCND---PLPINTSLAQNIAWLYPDSSEETV--LQVSWYVVSAS 132  
DB 108 WDCNLEEVAAQQLAPCND---PLPINTSLAQNIAWLYPDSSEETV--LQVSWYVVSAS 166  
QY 133 LGFMKGTKLDQ---FANQWAEPLANIANYRNKRVGCAHKICPAQNMVVSQVY 182  
DB 167 -GEEKGNE-DQNRGFYARRNYFGMANMAKGTYVGCSTYIMGDDGSALFTCLY 218

RESULT 15

Q9BI08\_9BILA PRELIMINARY; PRT; 248 AA.  
AC Q9BI08;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)  
DE Activation associated secreted protein-like protein (Fragment).  
OS Cooperia punctata.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongyloida;  
OC Trichostrongyloidea; Cooperiidae; Cooperia.  
NCBI\_Taxid=96640;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=CP-ASPar-7;  
RX PubMed=12443689; DOI=10.1016/S0034-5288(02)00125-X;  
RA Yacenda A.P., Eyeker M., Vieira-Bressan M.C., De Vries E.;  
RT "A family of activation associated secreted protein (ASP) homologues  
of Cooperia punctata."  
RL Reg. Vet. Sci. 73:297-306(2002).  
DR EMBL; AF352698; AAK35183.1; -; mRNA.  
DR HSP; 005110; 10NX.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR InterPro; IPR001283; Allrgn\_V5/Tpkl.  
DR InterPro; IPR002413; V5\_allergen.  
DR Pfam; PF00188; SCP, 1.  
DR PRINTS; PR00838; VSALLERGEN.  
DR PRINTS; PR00837; VSTPLIKE.  
DR SMART; SM00198; SCP, 1.  
DR PROSITE; PS01009; CRISP\_1; 1.  
FT NON TER 1  
SQ SEQUENCE 248 AA; 27558 MW; A7EDA83B9BEF7481 CRC64;

Query Match 14.0%; Score 176; DB 2; Length 248;  
Best Local Similarity 25.5%; Pred. No. 1.4e-07;  
Matches 53; Conservative 27; Mismatches 100; Indels 28; Gaps 7;

QY 23 CPNLSQSQDSARQIFLDNDVRRNIALGGLINWTVNADAVILGPAQNMVYKVDNCNLE 82  
DB 49 CTLDNGMTDBARQVFLDKNEIRQLVARG-----AQNKTGLAPPAAMLKLRIDCDLE 102  
QY 83 EYAAQQLAFC--NDPLPINTSLAQNIAWLYF---KQSEETVLQOVSWYVVSASLQFM 136  
DB 103 AHVMEHVAKCKGSHSPDVLKRGQNI--WALTVPMLDKAEAKRSVDDWYFELTKYGIT 160  
QY 137 KCTKLDQFANQWAEPLANIANY-----RNKRVGCAHKICPAQNMVVSQVYSPKLAPN 190  
DB 161 ADNKIS-----IDNAAKTGHYSQVVMOKSNRLGCAVSCPEQRRLVYGCETWPGANTLR 214  
QY 191 EVIMQGRACVCDARPDSCDNLCDTR 218

DB 215 HLTYDIGBPCKD--EDCKSCSCRCSTQ 240

Search completed: March 3, 2006, 19:09:28  
Job time : 231 secs

***This Page Blank (uspto)***

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

- Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).
- Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (uspto)**

## **Protein Sequence Searches - February 2005**

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

***This Page Blank (uspto)***